The identification and characterization of interactions between the heterotrimeric G protein beta subunit and two downstream effectors

Erin Julie Friedman

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology.

Chapel Hill
2011

Approved by:

Dr. Alan Jones

Dr. Jeff Dangl

Dr. Corbin Jones

Dr. Joe Kieber

Dr. Brenda Temple

ABSTRACT

ERIN JULIE FRIEDMAN: The identification and characterization of interactions between the heterotrimeric G protein beta subunit and two downstream effectors (Under the direction of Alan Jones)

Cells have evolved mechanisms to sense and respond to signals from the external environment and from other cells. One such mechanism is the heterotrimeric guanine-nucleotide binding protein (G protein) signaling pathway. This suite of signaling components allows cells to perceive extracellular signals using a seven-pass transmembrane receptor and to effect change within the cell via the dissociation of the G protein Gα subunit and Gβγ heterodimer, which both regulate downstream interacting proteins. In my dissertation work, I identified and characterized several interactions between the Gβ subunit and its downstream effectors.

First, I utilized an evolutionary approach to predict non-overlapping binding interfaces on the Gβ protein surface. Gβ proteins bind to a diverse group of interacting proteins, but structural data for many protein complexes are not available. I mapped the conservation of the Gβ surface residues over time to identify eight regions of interest (ROIs) that I hypothesized composed novel binding interfaces. By analyzing the evolutionary history of known Gβ effectors, I determined that one such ROI is required for the activation of the effector phospholipase C β2.

I next utilized a screen in the model plant *Arabidopsis thaliana* to identify genes whose overexpression compensated for the loss of *Arabidopsis* Gβ (AGB1) and therefore

were in the same genetic pathway as AGB1. Aci-reductone dioxygenase 1 (ARD1) interacts both genetically and physically with AGB1. I showed that ARD1 operates as an enzyme in the methionine salvage pathway, and AGB1 stimulates the enzymatic activity of ARD1. I used an evolutionary comparison to identify three ARD1 interaction sites on AGB1, and showed that mutation of these sites abolished AGB1 stimulation of ARD1. Finally, I predicted and tested the structural mechanism of ARD1 basal and stimulated activities, showing that a conserved tryptophan residue near its active site is critical for both functions, and a regulatory helix is critical for the stimulation of ARD1 by AGB1. Together, my results identified ARD1 as a novel Gβ effector, characterized its function in *Arabidopsis* and its regulation by AGB1, and provided structural insight into the mechanism of the regulation of downstream effectors by Gβ.

ACKNOWLEDGEMENTS

I would first like to thank Alan for giving me a research and a teaching home while at UNC. He has challenged me, and as a result, I am a better scientist. Alan also has respected my desire to pursue a teaching career from the first day I joined his lab; I am thankful for the opportunities he has provided to me to further my teaching experience, especially by allowing me to contribute directly to his courses. I am also thankful to my committee members Corbin, Brenda, Jeff, and Joe for their support. Their combined scientific expertise and perspective allowed me to take my research to places that we wouldn't have gone otherwise. When you work on a project so intimately, it is hard to make objective decisions and observations; each of my committee members provided this perspective for Alan and me when we would sometimes overlook the obvious. I would also like to thank Brenda and Corbin for their collaboration and friendship, inside and outside of the lab (and to Corbin for twelve consecutive hours watching *Lord of the Rings*).

The Jones lab has constantly changed during my time as a graduate student, but the common thread is that I have always been surrounded by helpful and encouraging colleagues. Being the only graduate student in a lab can be tough, but my colleagues were there every step of the way. They contributed to my work not only scientifically, but socially as well. I would like to specifically thank Helen Wang for initiating the SGB project (and then handing it over to me), and Jing Yang, Kun Jiang, Nguyen Phan, Jan

Jones, and Daisuke Urano for assisting me with various aspects of this project. I thank Yash Mudgil and May Christian for their continuing support and friendship. I always look forward to discussing the ins and outs of teaching with Tyrell Carr. I have been fortunate to mentor several fantastic undergraduates; Casey Kolb, Will Hannah, and Arwen Frick-Cheng not only contributed to my projects scientifically, but also became friends. Finally, I would like to thank Phil Taylor. Although we only overlapped in the Jones lab for a year (yes, Phil, I forgive you for leaving), he made that first year bearable. I always look forward to our visits with Phil, Amanda, and their two boys Jack and Hugh, and I am thankful that our families were brought together by this process. Maybe we will make it to that shuttle launch after all!

I have made lifelong friends in several of my Biology Department classmates. Sarah Taylor and I met at the airport during UNC recruitment weekend and have stayed close ever since. We interviewed, rotated, commiserated, and navigated motherhood together, and I am thrilled that she, Danny, and Mia are still willing to attempt weekly movie nights with us. We enjoy indulging our nerdy (gaming) side as often as possible with Jacob and Jessica, Dan and Kim, and Matt and Kelly. Mindy has kept me sane and grounded and served as my go-to stress relief when I just needed somebody to listen. I love that Kim, Mindy, and the other two "first years," Ben and Nathan, are always willing (even eager) to drink the Manny at Passover. I would also like to thank Emily and Jan for their continued friendship and support.

All of my friends from before graduate school (a.k.a. my previous life) have also contributed to my development and I wouldn't be here without them. A big shout out goes to The Guys from Dallas (yes, all of you!), the Trinity Family, and all of my

childhood friends (whose title is a misnomer – a more appropriate term is lifelong friends!). I would especially like to thank Meg (and Lucas) for the lifelong support, friendship, love, and for being awesome Godparents to Gabe. Jesse followed fairly closely in my footsteps while we grew up, but he beat me to both the PhD and the faculty job – hopefully I'm not too far behind the latter! Tommy and Erica were my lifeblood in college (and introduced me to Mike), and I am pleased to be able to discuss the joys and frustrations of teaching with Tommy and to navigate the process of graduate school with Erica (one more year at UNM, girlie!).

Many of my teachers molded me into the person and student that I am. I honestly would thank every one of them if I could, but would especially like to remember Larry Lynch for inspiring me in AP Biology and throughout high school academic competitions and to thank Dr. Jim Shinkle and Dr. Bob Blystone for mentoring me while at Trinity. Jim told my dad when I was a sophomore that he knew what I would do in life; I hope this was it! I wouldn't be at UNC if it weren't for the guidance of Dr. David Russell while I worked in his lab as a research technician at UT Southwestern in Dallas. David, a UNC Chemistry PhD, put UNC on my radar and encouraged me to pursue graduate studies.

Finally, I would like to thank my family. If you know me, you know that family lines are very vague and that we welcome everyone to join our group. I have been blessed by a close-knit network of family, in-laws, and friends who are always there for me even though we are so spread out geographically. In a nice change of events, we had family move here last summer. Sara, Ben, and Stella have been awesome supporters and provide a nearly weekly distraction for us; we love having you guys in town and can't wait to go

camping this spring! My parents Lynnette and Steve have always inspired me to do my best and to never stop learning, both by words and by example (as Mom taught third grade and Dad teaches college Biology). I not only grew up valuing education, but tooling around the lab at Franklin College. With Dad's help, I was entering (and winning) science fairs as early as kindergarten, and Mom let me help grade her students' homework assignments as soon as I was older than they were. Both have provided extraordinary support throughout my life and graduate career; I wouldn't have made it through grad school without you guys! My brother Adam always joked that he would never be a teacher or a scientist; I am both, and I chalk that up to typical older/younger sibling dynamics. Speaking of brothers, I wouldn't be who I am today without Adam and Rick. They are both very different from me, and yet have molded me and taught me to be strong, independent, and to love. Our times together are too infrequent, but always special (and there is always the possibility that I will become the most overqualified writer's assistant of all time if Rick has his way!). Finally, I want to remember Grandpa; he always wanted to know what I was studying and how school was going, and he was such a champion of my graduate studies. I will never forget our last ever conversation, as I stood waiting for the bus in the dark after a long day in the lab, as we discussed which lab I would eventually join and laughed over the phone together about everything going on in the world.

Finally, I want to thank Mike for his love, support, and willingness to move to NC in the first place. Even though we both wanted to pursue PhDs, I got the chance to go to grad school and Mike fully supported that decision; we packed up our TX home and he started life in a new state and in a new career to support my dream. He has not only been

supportive emotionally, but also scientifically; he worked for three summers in UNC biology labs, both for Alan and Corbin, and is as engrained in the department as I am. He has helped with numerous aspects of my research (from coding to plating seeds), and he has always encouraged me to stick with it even when all I wanted to do was quit. He has been an amazing husband and a wonderful father to our son, Gabriel. Speaking of Gabe, you are the best decision that your dad and I have ever made. You have made every day a complete joy; you bring a smile to my face that makes me forget even the worst day, and for that I can't thank you enough. I can't wait for the three of us to embark on our next adventure together!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AC2	Adenylyl cyclase 2
ACC	1-aminocycloprane-1-carboxylate
ACO	ACC oxidase
ACS	ACC synthase
amiRNA	Artificial microRNA
ARD	Aci-reductone dioxygenase
BiFC	Bimolecular fluorescence complementation
BLAST	Basic local alignment search tool
DMEM	Dulbecco's Modified Eagles Medium
E1	E1 enolase/phosphatase
ET	Evolutionary trace
G protein	Guanine nucleotide binding protein
GPCR	G protein coupled receptor
GRK	G-protein coupled receptor kinase
ICP-MS	Inductively coupled plasma mass spectrometry
MSA	Multiple sequence alignment
MS	Murashige & Skoog
MTA	S-methylthioadenosine
MTK	MTA kinase
MTN	MTA nucleosidase
PLC-β2	Phospholipase C beta 2
RGS9	Regulator of G-protein signaling 9
ROI	Region of interest

RT-PCR	Reverse transcriptase PCR
RSA	Relative solvent accessibility
SAM	S-adenosylmethionine
SAMS	SAM synthase
SPDS	Spermidine synthase
SPMS	Spermine synthase
SGB	Suppressor of G beta
Y3H	Yeast-three hybrid

CHAPTER 1

Introducing Heterotrimeric G Proteins

How do cells communicate with one another and sense and respond to their environments? Cells of eukaryotic organisms utilize a suite of signaling components that, together, form heterotrimeric guanine nucleotide binding protein (G protein) complexes. In the classical G protein signaling model, this complex is formed by an inactive seven-pass transmembrane G protein coupled receptor (GPCR) that is bound on the cytosolic surface of the plasma membrane to an inactive heterotrimeric G protein (Figure 1.1). This G protein consists of an inactive GDP-bound Gα subunit, a Gβ subunit, and a Gγ subunit. Both Gα and Gγ are tethered to the plasma membrane via covalent lipid modifications.

G protein signaling begins when a stimulus activates the GPCR, causing a conformational change in the cytoplasmic domains of the receptor. The guanine-nucleotide exchange factor activity of the GPCR induces a conformational change in the bound Gα-GDP, allowing Gα to release GDP. Because GTP is in excess in the cell, Gα readily binds GTP, activating the molecule. This activation allows Gα to dissociate from the Gβγ heterodimer (which remain tightly bound throughout the signaling process and require one another for protein stability), and exposes the shared binding surfaces. These surfaces, along with other solvent-exposed surfaces on the proteins, serve as sites of protein-protein interactions between each subunit and its downstream effectors (E1 and E2 in Figure 1.1). By binding to downstream effectors, Gα and Gβγ propagate signaling cascades. This signaling is terminated when the intrinsic GTPase activity of Gα hydrolyzes GTP to GDP and the inactive heterotrimer reforms (Tuteja, 2009; Tesmer, 2010). In addition to the Gα GTPase, accessory proteins such as those belonging to the regulator of G-protein signaling (RGS) family can accelerate this hydrolysis (Siderovski et al., 1999).

G protein complexity and evolution

Of course, G protein signaling is not as simple as this ideal paradigm suggests. Although the theme remains constant, many variations of this signaling regime have been identified. For instance, mammalian systems contain 23 Gα subunits, five Gβs, and twelve Gγs in addition to hundreds of GPCRs (Jones and Assmann, 2004). Dozens of downstream effectors of both Gβγ and Gα have been identified in addition to a number of scaffolding and regulatory molecules. The suite of components in any given cell is a product of the individual organism and the specific cell type. However, it has been shown that many components of the heterotrimer can bind promiscuously, such that several different combinations of heterotrimer subunit proteins or Gβγ dimers are possible (Clapham and Neer, 1997). These different subunit combinations enlarge the signaling capabilities of this complex while also providing signaling specificity.

G protein components have been grouped into families based on their known function, sequence homology, and evolutionary history. Gα proteins can be grouped into four main classes (G(io), G(q), G(s), and G(12)), while Gβ proteins can be grouped into two classes (Gβ1-like and Gβ5-like). For both molecules, the evolution that followed rounds of duplication was constrained in part by residues on the molecules' surfaces that are involved in protein-protein interactions with effectors, regulators, or other G protein components (Friedman et al., 2009; Temple et al., 2010). The primordial G protein components duplicated and diverged throughout evolutionary history, although the plant G protein components contain the most characteristics that are similar to the common ancestor (Temple and Jones, 2007).

Mammalian Gβ effectors and accessory proteins

Activated G protein subunits positively and negatively regulate downstream effector activity. Although researchers once believed that the sole functions of Gβ was to inhibit Gα signaling and to localize Gα to the membrane (Milligan and Kostenis, 2006), it is now clear that Gβ regulates downstream signaling cascades via interactions with a variety of effectors and accessory proteins. Gβ proteins have been studied extensively in mammalian systems; multiple mammalian effectors have already been identified. Gβ stimulates G protein-gated inwardly rectifying potassium channels (GIRKs), which regulate membrane potential in a cell-type dependent manner. Gβ activates phospholipase C β, ultimately resulting in the release of Ca^{++} into the cytoplasm. In neurons, Gβ inhibits the activity of N-type Ca^{++} channels, affecting the release of neurotransmitters (Mirshahi et al., 2002). Mammalian accessory proteins that interact with Gβ have also been identified. Retinal phosducin regulates G protein subunit availability, possibly by assisting in the proper localization Gβ or by sequestering Gβ in the cytoplasm; this regulation facilitates the adaptation to light (Gaudet et al., 1999). Additional proteins with similarity to phosducin are expressed more ubiquitously than retinal rhodopsin and may play function as signaling molecules or localization chaperones (Gaudet et al., 1999; Krispel et al., 2007). The RGS R7 family of proteins binds preferentially to Gβ5-like proteins via a G-gamma-like (GGL) domain on RGS. The RGS GGL domain binding to Gβ precludes the binding of Gγ, resulting in a Gβ-RGS heterodimer instead the traditional Gβγ heterodimer. This is the only known example whereby Gβ is stabilized by a protein other than Gγ (Snow et al., 1998; Cheever et al., 2008).

The three-dimensional structure of Gβ provides insights into its function and regulation of downstream effectors. Gβ is formed by an N-terminal alpha helix of approximately twenty residues, which interacts with Gγ, and a C-terminal seven-bladed beta-propeller that contains seven WD repeat sequences (Sondek et al., 1996). When inactive, the "top" surface of this beta propeller structure is occupied by the binding of Gα; upon activation, no conformational change occurs on the Gβ molecule, but this binding surface is exposed and is highly utilized to form a portion of the binding interfaces between Gβ and its downstream effectors. Thus, the interactions between Gβ and its effectors are regulated by its binding site availability. Within this binding region, several key residues termed "hot spots" form critical contacts between Gβ and its effectors (Smrcka, 2008).

G proteins in plants

The G protein signaling regime in the model plant species *Arabidopsis thaliana* is simplified when compared to those of mammalian species. *Arabidopsis* contains one gene encoding Gα (GPA1), one encoding Gβ (AGB1), and at least two genes encoding Gγ (AGG1 and AGG2) (Jones and Assmann, 2004). Although only two Gγ genes have been identified, it is now clear that additional genes may exist because the double knockout of Gγ1 and Gγ2 does not recapitulate the Gβ-null phenotype, nor do the localization patterns of Gγ1 and Gγ2 completely overlap with Gβ (Trusov et al., 2008). Because of this relative simplicity, *Arabidopsis* is an ideal system in which to study G protein signaling. With the few number of possible heterotrimer combinations and because null mutations do not confer lethality, genetic studies can be performed on plants lacking complete

portions of the G protein system. However, in plants, no downstream effectors of Gβ have been previously identified (Jones and Assmann, 2004).

G protein components in plants mediate a multitude of downstream pathways. Plant G proteins modulate cell growth and division (Ullah et al., 2001), seed germination (Ullah et al., 2002), ion channels (Wang et al., 2001), responses to stresses such as ozone (Joo et al., 2005) and pathogen infection (Llorente et al., 2005; Trusov et al., 2006), and responses to hormones (Perfus-Barbeoch et al., 2004; Steffens and Sauter, 2010).

Of the previously characterized mutations in *Arabidopsis* G protein subunits, those lacking a functional AGB1 gene (*agb1-2*) display a strong, wide array of phenotypes (Ullah et al., 2003). While the ectopic overexpression of AGB1 induces cell division, *agb1* mutants display reduced cell division (Lease et al., 2001). *agb1-2* hypocotyls display decreased adaxial but increased circumferential cell division, while roots display increased cell division. Two-day-old etiolated *agb1-2* seedlings have a shortened hypocotyl and opened apical hook when compared to wildtype seedlings, and *agb1-2* plants have excessive lateral root formation that is induced by auxin. Ten-day-old light-grown *agb1-2* cotyledons are larger and rounder than their wildtype counterparts, and the siliques of mature plants are blunt and short (Ullah et al., 2003). Two-week-old *agb1-2* plants are hypersensitive to infection by the necrotrophic fungus *Fusarium oxysporum* (Trusov et al., 2006). This wide array of phenotypes suggests that AGB1 mediates multiple signaling pathways, and it therefore likely interacts with a variety of downstream effectors.

Summary

Based on the current knowledge gained from both plants and higher eukaryotic systems, we can utilize the study of *Arabidopsis* G protein signaling in three ways. First, as we have done in the past, we can utilize the knowledge from higher eukaryotes to make predictions about G protein signaling in plants. For example, the knowledge of mammalian GPCRs has been utilized to screen plant systems for the presence of similar putative GPCRs (Moriyama et al., 2006). Second, we can identify plant-specific functions of G protein signaling, such as responses to plant-specific hormones like ethylene (Steffens and Sauter, 2010). Finally, we can apply findings from the plant G protein system to mammalian and human systems. By identifying novel effectors in plants that are also maintained in mammalian organisms, we can determine whether these proteins are also mammalian G protein effectors. This is critical, because due to the roles played in cell growth and proliferation, G protein signaling has been implicated in a number of diseases and cancers (Lappano and Maggiolini, 2011). GPCRs (and thus their downstream effectors) are targeted by ~40% of the pharmaceutical drugs currently on the market, and the targeting of GPCRs or even directly targeting G protein subunits is an active field of biomedical and pharmacological research (Siehler, 2008; Smrcka et al., 2008).

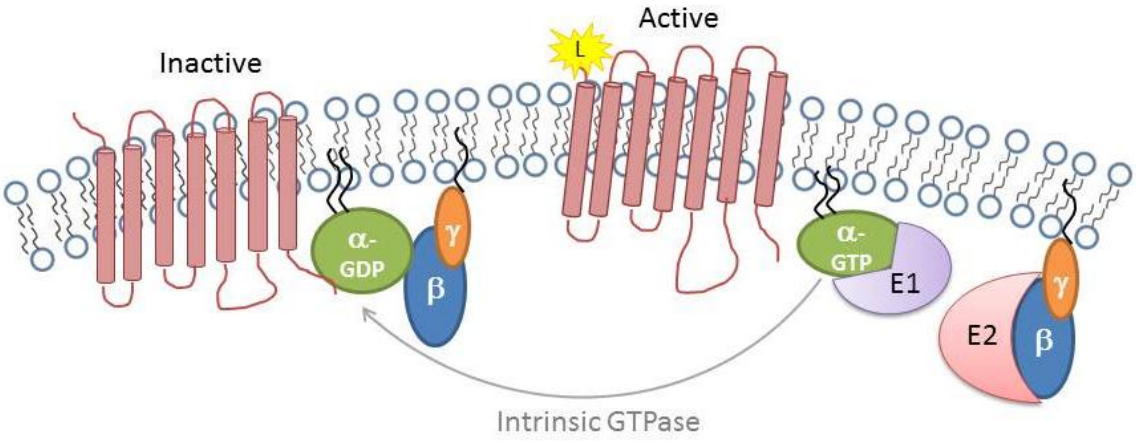


Figure 1.1. Canonical heterotrimeric G protein signaling. Inactive and active G protein complexes are shown as discussed in the text.

References

Cheever, M.L., Snyder, J.T., Gershburg, S., Siderovski, D.P., Harden, T.K., and Sondek, J. (2008). Crystal structure of the multifunctional Gbeta5-RGS9 complex. Nat Struct Mol Biol **15,** 155-162.

Clapham, D.E., and Neer, E.J. (1997). G protein beta gamma subunits. Annu Rev Pharmacol Toxicol **37,** 167-203.

Friedman, E.J., Temple, B.R., Hicks, S.N., Sondek, J., Jones, C.D., and Jones, A.M. (2009). Prediction of protein-protein interfaces on G-protein beta subunits reveals a novel phospholipase C beta2 binding domain. J Mol Biol **392,** 1044-1054.

Gaudet, R., Savage, J.R., McLaughlin, J.N., Willardson, B.M., and Sigler, P.B. (1999). A molecular mechanism for the phosphorylation-dependent regulation of heterotrimeric G proteins by phosducin. Mol Cell **3,** 649-660.

Jones, A.M., and Assmann, S.M. (2004). Plants: the latest model system for G-protein research. EMBO Rep **5,** 572-578.

Joo, J.H., Wang, S., Chen, J.G., Jones, A.M., and Fedoroff, N.V. (2005). Different Signaling and Cell Death Roles of Heterotrimeric G Protein {alpha} and {beta} Subunits in the Arabidopsis Oxidative Stress Response to Ozone. Plant Cell **17,** 957-970.

Krispel, C.M., Sokolov, M., Chen, Y.M., Song, H., Herrmann, R., Arshavsky, V.Y., and Burns, M.E. (2007). Phosducin regulates the expression of transducin betagamma subunits in rod photoreceptors and does not contribute to phototransduction adaptation. J Gen Physiol **130,** 303-312.

Lappano, R., and Maggiolini, M. (2011). G protein-coupled receptors: novel targets for drug discovery in cancer. Nat Rev Drug Discov **10,** 47-60.

Lease, K.A., Wen, J., Li, J., Doke, J.T., Liscum, E., and Walker, J.C. (2001). A mutant Arabidopsis heterotrimeric G-protein beta subunit affects leaf, flower, and fruit development. Plant Cell **13,** 2631-2641.

Llorente, F., Alonso-Blanco, C., Sanchez-Rodriguez, C., Jorda, L., and Molina, A. (2005). ERECTA receptor-like kinase and heterotrimeric G protein from Arabidopsis are required for resistance to the necrotrophic fungus Plectosphaerella cucumerina. Plant J **43,** 165-180.

Milligan, G., and Kostenis, E. (2006). Heterotrimeric G-proteins: a short history. Br J Pharmacol **147 Suppl 1,** S46-55.

Mirshahi, T., Mittal, V., Zhang, H., Linder, M.E., and Logothetis, D.E. (2002). Distinct sites on G protein beta gamma subunits regulate different effector functions. J Biol Chem **277,** 36345-36350.

Moriyama, E.N., Strope, P.K., Opiyo, S.O., Chen, Z., and Jones, A.M. (2006). Mining the Arabidopsis thaliana genome for highly-divergent seven transmembrane receptors. Genome Biol **7,** R96.

Perfus-Barbeoch, L., Jones, A.M., and Assmann, S.M. (2004). Plant heterotrimeric G protein function: insights from Arabidopsis and rice mutants. Curr Opin Plant Biol **7,** 719-731.

Siderovski, D.P., Strockbine, B., and Behe, C.I. (1999). Whither goest the RGS proteins? Crit Rev Biochem Mol Biol **34,** 215-251.

Siehler, S. (2008). Cell-based assays in GPCR drug discovery. Biotechnol J. **3,** 471-483.

Smrcka, A.V. (2008). G protein betagamma subunits: central mediators of G protein-coupled receptor signaling. Cell Mol Life Sci **65,** 2191-2214.

Smrcka, A.V., Lehmann, D.M., and Dessal, A.L. (2008). G protein betagamma subunits as targets for small molecule therapeutic development. Comb Chem High Throughput Screen **11,** 382-395.

Snow, B.E., Krumins, A.M., Brothers, G.M., Lee, S.F., Wall, M.A., Chung, S., Mangion, J., Arya, S., Gilman, A.G., and Siderovski, D.P. (1998). A G protein gamma subunit-like domain shared between RGS11 and other RGS proteins specifies binding to Gbeta5 subunits. Proc Natl Acad Sci U S A **95,** 13307-13312.

Sondek, J., Bohm, A., Lambright, D.G., Hamm, H.E., and Sigler, P.B. (1996). Crystal structure of a G-protein beta gamma dimer at 2.1A resolution. Nature **379,** 369-374.

Steffens, B., and Sauter, M. (2010). G proteins as regulators in ethylene-mediated hypoxia signaling. Plant Signal Behav **5,** 375-378.

Temple, B.R., and Jones, A.M. (2007). The plant heterotrimeric G-protein complex. Annu Rev Plant Biol **58,** 249-266.

Temple, B.R., Jones, C.D., and Jones, A.M. (2010). Evolution of a signaling nexus constrained by protein interfaces and conformational States. PLoS Comput Biol **6,** e1000962.

Tesmer, J.J. (2010). The quest to understand heterotrimeric G protein signaling. Nat Struct Mol Biol **17,** 650-652.

Trusov, Y., Zhang, W., Assmann, S.M., and Botella, J.R. (2008). Ggamma1 + Ggamma2 not equal to Gbeta: heterotrimeric G protein Ggamma-deficient mutants do not recapitulate all phenotypes of Gbeta-deficient mutants. Plant Physiol **147,** 636-649.

Trusov, Y., Rookes, J.E., Chakravorty, D., Armour, D., Schenk, P.M., and Botella, J.R. (2006). Heterotrimeric G proteins facilitate Arabidopsis resistance to necrotrophic pathogens and are involved in jasmonate signaling. Plant Physiol **140,** 210-220.

Tuteja, N. (2009). Signaling through G protein coupled receptors. Plant Signal Behav **4,** 942-947.

Ullah, H., Chen, J.G., Wang, S., and Jones, A.M. (2002). Role of a heterotrimeric G protein in regulation of Arabidopsis seed germination. Plant Physiol **129,** 897-907.

Ullah, H., Chen, J.G., Young, J.C., Im, K.H., Sussman, M.R., and Jones, A.M. (2001). Modulation of cell proliferation by heterotrimeric G protein in Arabidopsis. Science **292,** 2066-2069.

Ullah, H., Chen, J.G., Temple, B., Boyes, D.C., Alonso, J.M., Davis, K.R., Ecker, J.R., and Jones, A.M. (2003). The beta-subunit of the Arabidopsis G protein negatively regulates auxin-induced cell division and affects multiple developmental processes. Plant Cell **15,** 393-409.

Wang, X.Q., Ullah, H., Jones, A.M., and Assmann, S.M. (2001). G protein regulation of ion channels and abscisic acid signaling in Arabidopsis guard cells. Science **292,** 2070-2072.

CHAPTER 2

Prediction of Protein-Protein Interfaces on G-Protein β Subunits Reveals a Novel Phospholipase C β2 Binding Domain

Preface

For the second chapter, I have included my first-author paper that was published in 2009 in the Journal of Molecular Biology and entitled "Prediction of protein-protein interfaces on G-protein β Subunits Reveals a Novel Phospholipase C β2 Binding Domain." The co-authors on this paper were Brenda R.S. Temple, Stephanie N. Hicks, John Sondek, Corbin D. Jones, and Alan M. Jones. Brenda Temple assisted me with the evolutionary analysis, and together we experimented with different ways of examining the evolution of interfaces before settling on the version published here. Stephanie Hicks, a research assistant professor in the lab of John Sondek, performed the mutagenesis and activation experiments. Corbin Jones provided evolutionary analysis and critique. This work was performed under the direction of Alan Jones.

Abstract

Gβ subunits from heterotrimeric G-proteins directly bind diverse proteins, including effectors and regulators, to modulate a wide array of signaling cascades. These numerous interactions constrained the evolution of the molecular surface of Gβ. Though mammals contain five Gβ genes comprising two classes (Gβ1-like and Gβ5-like), plants and fungi have a single ortholog and organisms such as *Caenorhabditis elegans* and *Drosophila melanogaster* contain one copy from each class. A limited number of crystal structures of complexes containing Gβ subunits and complementary biochemical data highlight specific sites within Gβs needed for protein interactions. It is difficult to determine from these interaction sites what, if any, additional regions of the Gβ molecular surface comprise interaction interfaces essential to Gβ's role as a nexus in numerous signaling cascades. We used a comparative evolutionary approach to identify five known and eight previously-unknown putative interfaces on the surface of Gβ. We show that one such novel interface occurs between Gβ and phospholipase C β2 (PLC-β2), a mammalian Gβ interacting protein. Substitutions of residues within this Gβ-PLC-β2 interface reduce the activation of PLC-β2 by Gβ1, confirming that our *de novo* comparative evolutionary approach predicts previously unknown Gβ-protein interfaces. Similarly, we hypothesize the seven remaining untested novel regions contribute to putative interfaces for other Gβ interacting proteins. Finally, this comparative evolutionary approach is suitable for application to any protein involved in a significant number of protein-protein interactions.

Introduction

Gene duplication is a fundamental source of genetic and phenotypic novelty (Ohno, 1970). After duplication, one of the new paralogs is freed from functional constraint, enabling it to evolve new functions. For genes encoding proteins that interact with other proteins, this process often liberates one copy to develop a new set of interactions. This is particularly true for signaling molecules that are used by organisms to communicate between cells and to perceive their environment. For example, evolution of increasingly complex organisms correlates with the enormous diversification of heterotrimeric guanine nucleotide-binding protein (G-protein) signaling complexes and cell surface G-protein coupled receptors. The G-protein complex consists of a heterotrimer comprised of Gα, Gβ, and Gγ subunits. Upon activation by cell surface receptors, the complex dissociates into a free Gα subunit and a Gβγ dimer, both of which bind to and signal through other proteins. Signaling typically terminates when the heterotrimer reforms (Jones and Assmann, 2004; McCudden et al., 2005). *Saccharomyces cerevisiae* have two Gα but a single Gβ and a single Gγ gene; mammalian genomes encode sixteen Gα, five Gβ, and twelve Gγ genes (Jones and Assmann, 2004). This complex array of subunit combinations allows for diverse signaling possibilities. It was previously thought that Gα was the primary signaling molecule in mammals while the sole function of Gβ was to inhibit Gα signaling and to provide for its membrane localization (Milligan and Kostenis, 2006). It is now clear that Gβ also modulates downstream targets; a subset of signaling pathways is uniquely regulated by the Gβ subunit (Cabrera-Vera et al., 2003).

Several proteins that bind Gβ have been identified in mammals. In addition to Gα and Gγ, interacting proteins such as the localization chaperone phosducin, G-protein-coupled receptor kinases (GRKs), phospholipase C β2 (PLC-β2), regulator of G-protein signaling 9 (RGS9), calcium channels, potassium channels, and adenylyl cyclase 2 (AC2) bind Gβ proteins. Mammalian Gβ interacting proteins arose differentially over evolutionary time as not all eukaryotes contain homologs to all known interactors (Jones and Assmann, 2004). For example, Gα and phosducin are present in all eukaryotes beginning with plants. Canonical RGS9, PLC-β2, and GRK2 originated and were maintained in metazoans at least by the time of the formation of annelids, since *Caenorhabditis elegans* contains bona fide RGS9, PLC-β2, and GRK2 orthologs (Miller et al., 1999; van der Linden et al., 2001; Fukuto et al., 2004). As the Gβ subunit acquired additional binding partners throughout evolution, new Gβ-protein interfaces likely evolved to accommodate these interactions. These new interfaces may have partially overlapped with existing interfaces since, for example, the interface from many Gβ interactors overlaps with the Gα-Gβ interface (Ford et al., 1998). However, these interfaces may have also utilized regions on the Gβ molecular surface that previously had no associated function.

Two major experimental approaches, structural and biochemical studies, have characterized some binding interfaces between Gβ and interacting proteins. Currently, there are four crystal structures of mammalian Gβ subunits in complex with signaling proteins: Gβ1γ1-Gα (Lambright et al., 1996), Gβ1γ2-GRK2 (Lodowski et al., 2003), Gβ1γ1-phosducin (Loew et al., 1998), and Gβ5-RGS9 (Cheever et al., 2008). These structures provide a three-dimensional view of where proteins interact, but provide

limited information regarding the importance of individual contacts. Additionally, structures of Gβ in complex with peptides (Davis et al., 2005; Johnston et al., 2008) provide partial information on physiologically relevant interfaces within Gβ. Biochemical studies include both targeted mutational studies (Ford et al., 1998; Li et al., 1998; Panchenko et al., 1998) and targeted domain-swapping or peptide binding experiments (Yoshikawa et al., 2001; Bonacci et al., 2005; Myung et al., 2006). Mutational studies directed by the structural studies are limited by the number of available structures. Interpretation of these studies is complicated by overlap between the implicated binding regions. When crystal structures of a particular Gβ complex are not available, mutations are frequently targeted to an interface identified in a solved Gβ complex structure. For example, Ford *et al.* (Ford et al., 1998) created several mutations in the Gα binding interface to show that this interface was utilized in part by five other interacting proteins. This type of approach does not, however, elucidate binding sites, or even those regions of the binding interface, that are not shared with the Gα interface. To locate these alternative sites, groups such as Panchenko *et al.* (Panchenko et al., 1998) performed mutational analyses outside of known binding areas. Although this study identified mutant regions of Gβ with reduced ability to activate PLC-β2, much refinement remains necessary to identify individual critical interaction sites. Moreover, this study left many regions of the Gβ surface unexplored.

Bioinformatic analyses have been developed to identify functional residues including those composing conserved patches on the surface of a protein that may function as a binding interface. Three such analyses include Evolutionary Trace (ET) (Lichtarge et al., 1996a), DIVERGE (Gu and Vander Velden, 2002; Zheng et al., 2007),

and ConSurf (Armon et al., 2001). All three rely on multiple sequence alignments (MSAs) and phylogenetic trees to identify structurally-clustered functional residues. ET has been applied to Gβ and the two predicted interfaces correlated with those of Gα and Gγ (Lichtarge et al., 1996b). To date, none of these three methods have predicted previously uncharacterized Gβ interfaces.

We took advantage of both the rich source of divergent Gβ subunit sequences and the differential evolutionary emergences of known Gβ interactors to understand not only how to predict novel interaction interfaces on the Gβ surface in the absence of crystallized complexes, but also how these new interactions gave rise to new binding surfaces. These results can be utilized for more targeted and informed biochemical studies. Based on the hypothesis that the acquisition of mammalian-like sequence identity on the surface of Gβ reflects the utilization of new binding interfaces, we applied a suite of bioinformatic and phylogenetic techniques to follow the shift in patterns of amino acid conservation in Gβs, concentrating on changes between distinct points in the evolutionary history of the Gβ subunit represented by five reference species. By placing this conservation in a structural context, we predicted regions of interest (ROI) that are comprised of adjacent surface residues that simultaneously evolved to residues conserved with mammals. We identified a novel PLC-β2 interface by demonstrating that at least one ROI which became conserved in *C. elegans* is involved in PLC-β2 activation by Gβ. Similarly, we propose that the remaining ROIs also compose at least portions of binding interfaces.

Results and Discussion

Gβ proteins fall into two major classes: Gβ1-like and Gβ5-like.

Since most extant plant and fungal species have a single Gβ, while nematodes and later metazoans have at least two Gβ subunits, the first Gβ gene duplication occurred between the splitting of fungi and *C. elegans* from the mammalian lineage. To compare pre-duplication plant and fungal Gβ sequences to extant post-duplication Gβ sequences from metazoans, ancestral sequences to plants and to fungi were reconstructed from MSAs (Figs. S1 and S2, respectively) of extant plant and fungal Gβ sequences using the Bayesian ancestral reconstruction implemented in MrBayes (see Methods). These inferred pre-duplication ancestors were important as they, and not individual plant or fungal species, reflect sequence constraints common to all plants or fungi and, therefore, more closely reflect the predecessor of all extant post-duplication Gβ genes.

An MSA was created containing Gβ sequences from the plant ancestor, fungal ancestor, and extant *C. elegans, Drosophila melanogaster*, and human Gβ sequences (Figure 2.S3) for comparison of pre- and post-duplication Gβ genes. This MSA was used to generate a Bayesian phylogenetic tree (Figure 2.S4) that elucidates the evolutionary relationship between the different Gβ proteins. Human Gβ1, Gβ2, Gβ3, and Gβ4 sequences formed a monophyletic Gβ1-like clade also containing *D. melanogaster* and *C. elegans* Gβ1 sequences. Human Gβ5 formed a Gβ5-like monophyletic clade containing *D. melanogaster* Gβ5 and *C. elegans* sequences Gβ2. Both plant and fungal ancestors were outside of these two clades.

Based on the phylogenetic data (Figure 2.S4) and corroborating previous observations that Gβ5 often behaves differently than the other four mammalian Gβ proteins (Cabrera-Vera et al., 2003), metazoan Gβ proteins can be grouped into two major classes, Gβ1-like and Gβ5-like. The initial Gβ gene duplication between fungi and

C. elegans gave rise to the Gβ5-like family since *C. elegans* is the earliest species examined to contain a Gβ5-like protein. The ancestral plant and fungal Gβ proteins each contain characteristics of both major classes, but neither strictly belongs to either class. Finally, the phylogeny revealed that after duplication, the Gβ5-like genes diverged from the ancestor more than the Gβ1-like genes while the Gβ1-like genes maintained more ancestral characteristics.

Interaction Interfaces Identified from Gβ Complex Structures Co-evolved with Interactors.

Complexes containing mammalian Gβ1-Gα, Gβ1-GRK2, Gβ1-phosducin, and Gβ5-RGS9 are shown in Figure 2.1 (Gβ in spheres, interacting protein in ribbons). While these structures provide a three-dimensional view of Gβ-protein complexes, the interfaces between the proteins must still be defined (Jones and Thornton, 1996). To define sites of interaction, we calculated the solvent accessibility of each residue in the four crystal structures using Naccess (Hubbard and Thornton, 1993). Those Gβ residues whose side chains have lower relative solvent accessibility by more than five percent between the monomer and the protein complex, including those that hydrogen bond with the binding partner were defined to comprise the interface and were noted on the structure (see Methods) (Figure 2.1). These four Gβ-interacting proteins: 1) cover three major areas on the Gβ surface, 2) all partially overlap with each other, and 3) appeared at different times over Gβ evolution.

In order to examine the evolution of these known Gβ-protein interfaces, we compared the patterns of amino acid conservation within these regions. For each interface residue in the MSA, comparisons were made between each of the four reference

sequences and the corresponding mammalian residue. For Gβ1 complexes (Gα, GRK2, and phosducin), the plant ancestor, fungal ancestor, *C. elegans* Gβ1, and *D. melanogaster* Gβ1 were compared to bovine Gβ1. Identical interface residues were mapped onto the bovine Gβ1 structure (1got.pdb). For the RGS9 complex, the plant ancestor, fungal ancestor, *C. elegans* Gβ5 and *D. melanogaster* Gβ5 were compared to mouse Gβ5 and identities were noted on the mouse Gβ5 structure (2pbi.pdb). As expected, conservation within known binding areas correlated with the known utilization of these regions (Figure 2.2). The following was observed in the four crystal structure complexes:

Gβ1-Gα. The Gβ-Gα interaction is the most ancient, thus it would be expected that the Gα-Gβ interface is well-conserved in all organisms studied. As expected, the Gα binding interface was highly conserved (77% - 100% identity) in all sequences analyzed (Figure 2.2a, top row) as illustrated by the predominantly orange coloration of the interface.

Gβ1-GRK2. The GRK2 interface (Figure 2.2a, second row) was completely conserved in *C. elegans and D. melanogaster*, correlating with the emergence of GRK2 between fungi and *C. elegans*. Though a large portion of the interface was also conserved in the plant (62%) and fungal (92%) ancestors, this can be attributed to the fact that the interface largely overlaps that of Gα.

Gβ1-phosducin. Plants contain several phosducin-like sequences (Blaauw et al., 2003) and our yeast-three-hybrid data suggest that at least one phosducin in *Arabidopsis thaliana* interacts with the *A. thaliana* Gβ (Figure 2.3). The phosducin interface is highly conserved in *C. elegans* and *D. melanogaster* (94% identity); however, it is only partly conserved in plant and fungal ancestors (55% and 64%, respectively). Mammalian

phosducin has two domains: a helical N-terminal domain and a C-terminal domain (see Figure 2.1 for locations of N- and C-terminal binding interfaces), both of which bind mammalian Gβ (Gaudet et al., 1996). The plant and fungal ancestors show high conservation in the N-terminal binding area (68% and 76%, respectively) but not in the C-terminal binding area (25% and 38%, respectively) (Figure 2.2a, third and fourth rows). Two possible explanations for these results are that plants and fungi bind only one phosducin domain but not the second or that the second domain binds in a species-specific manner.

Gβ5-RGS9. The RGS9 binding interface (Figure 2.2b) is only 42% identical in the plant ancestor and 50% in the fungal ancestor. Correlating with the genesis of the RGS9 protein, conservation in the binding area rises to 67% and 76% in *C. elegans* and *D. melanogaster*, respectively. Though this is markedly less than the 90% to 100% identity seen in corresponding Gβ1 binding areas, it is important to note the size of the RGS9 interface. Encompassing 103 residues, the RGS9 interface is over three times larger than the Gα (31 residues) and GRK2 (26 residues) interfaces and twice the size of the phosducin interface (47 residues). Thus, we expect the RGS9 interface could tolerate more substitutions than its smaller counterparts and that only those residues making energetically critical contacts in the interaction interface are conserved.

Newly Conserved Regions of Adjacent Surface Residues Arose Over Time and Likely Contribute to Gβ-protein Interfaces.

As conservation within known Gβ-protein interfaces correlated with the emergence of proteins utilizing these interfaces, we next used a similar method of

analysis to predict *novel* binding interfaces. This unbiased comparative evolutionary analysis was not restricted to predetermined binding interfaces and was performed on the entire molecule, but was otherwise similar to the methods used in the validation stage described in the previous section. Bovine Gβ1 was compared at each residue in the MSA to the plant ancestor, fungal ancestor, *C. elegans* Gβ1, and *D. melanogaster* Gβ1 (Figure 2.4a). Identical residues between the plant ancestor and bovine Gβ1 were located on the surface of Gβ1 and colored dark green. These 162 residues form region zero (ROI0) and represent the primordial function of the Gβ molecule, since the plant molecule is the most ancestral-like of the Gβ family. For all other reference species, conserved residues that continued to be conserved from the previous species were colored light green (indicating the persistence of an existing function) while newly conserved residues were colored dark green (indicating the emergence of potential novel function). Similarly, mouse Gβ5 was compared to the plant ancestor, fungal ancestor, *C. elegans* Gβ5, and *D. melanogaster* Gβ5 and conserved residues were mapped onto the structure (Figure 2.4b).

Due to its primordial nature, ROI0 contains residues conserved both for structural maintenance as well as for interaction surfaces. ROI0 contains the residues – *e.g.* W99, M101 K57, Y59, L117, D186, D228, and W332 – previously indicated as interacting with multiple effectors and which also made energetically critical contacts in various interaction interfaces (Ford et al., 1998; Davis et al., 2005). Subsequent non-ancestral ROIs in each organism (for both Gβ1 and Gβ5) were defined as clusters of three or more structurally adjacent (within 5 Å (Madabushi et al., 2002)) newly conserved surface residues and are encircled on the structure (Figure 2.4, for residue designations see Table 2.1). It is important to note that though all surface-exposed residues could form a part of

an interface, we chose a larger cluster size (containing three or more newly-fixed residues) in order to predict clusters that likely make significant energetic contributions to binding. We also ignored several large clusters near the Gγ binding area due to their likely involvement in the Gβ-Gγ interface.

Five of the thirteen ROIs (7, 10, 11, 12, and 13) could be explained by existing structural data. ROI7 lies in the phosducin binding region on Gβ1 (Fig 1). Portions of ROIs 10-13 localize to the RGS9 binding interface on Gβ5 (Fig 1). Additionally, a portion of ROI2 was shown to interact with Gα, AC2, calcium channels, potassium channels, and PLC-β2 in mutational studies (Ford et al., 1998). These data support the hypothesis that regions of adjacent surface residues that become conserved to the mammalian state at the same time are likely sites of Gβ-protein interactions. Similarly, we hypothesize that the remaining ROIs are binding interfaces with yet-to-be identified interacting proteins or additional, undescribed binding interfaces with previously identified interacting proteins.

Several ROIs Contain Gβ Residues That Activate PLC-β2.

To test our hypothesis that ROIs represent Gβ-protein interaction interfaces, we mutated residues likely to interact with PLC-β2. Despite the absence of a crystal structure of the Gβγ-PLC-β2 complex, several studies implicated multiple potential interfaces. Peptide-binding assays implicated a region at the base of Gβ's N-terminal helix (Yoshikawa et al., 2001; Bonacci et al., 2005). A patch of residues near this region was implicated as an interface when it was revealed that the C-terminal tail of Gγ was responsible for the activation of PLC-β2 (Myung et al., 2006). Therefore, ROIs 4, 6, and 7 were chosen for mutational studies due to their proximity to the general region

implicated in these peptide-binding and domain-swapping experiments. Additionally, all three regions became conserved between the fungi and *C. elegans* split, corresponding with the emergence of PLC-β proteins.

In order to evaluate the extent of PLC-β2 activation, COS-7 cells were transfected with wild-type and mutant forms of Gβ1 and wild-type Gγ2 in the presence or absence of PLC-β2. Equal expression of all Gβ1, Gγ2, and PLC-β2 constructs was confirmed by immunoblot analysis. Additionally, equivalent expression of Gγ2 was observed in the presence of all forms of Gβ1. The location of each mutation on the surface of Gβ1 is depicted in Figure 2.5. The construct containing the double mutation T65A_D322A, which is located outside of the predicted binding region and lacks sequence conservation with bovine Gβ1, activates PLC-β2 to the same extent as wild-type Gβγ. Conversely, two single mutations W99A and D228A abolish PLC-β2 activation as previously shown (Ford et al., 1998). Within the newly predicted binding areas, the mutations that most significantly reduced PLC-β2 activation were found in ROIs 4 and 6 (Figs. 4a and 5). R52S_F335S slightly reduced PLC-β2 activation while K127S_R129S_E130S nearly abolished PLC-β2 activation. In summary, when residues within the new binding areas predicted by our *de novo* analysis are mutated there is a concordant drop in activity; conversely, when residues we identified as not part of binding surfaces are mutated there is no change in activation. These results indicate that our evolutionary structural analysis predicts the non-overlapping portions of interaction interfaces and can be used to further dissect the PLC-β2 interface on Gβ1.

Molecular Evolution Can Be Used as a Tool to Predict Novel Binding Interfaces.

Conservation is a hallmark of a residue's functional importance (Zheng et al., 2007) and as the plant ancestor is most similar to the most ancient Gβ molecule, primordial Gβ function can be inferred to correlate with those residues which were conserved in the plant ancestor (ROI0) (Figure 2.4). These primordial functional regions comprise both conserved protein-binding interfaces as well as regions of the protein important for structural integrity. With each comparative step (*e.g*. between plants and fungi, fungi and *C. elegans*, etc.), new functional regions became conserved. These regions may represent sites of novel Gβ-protein interaction or new interaction sites with existing proteins (co-evolution). Lineage-specific protein interactions, however, will not be revealed by this type of analysis. Lineage-specific interactions would be indicated by conservation of a residue within a particular organismal group (*e.g.* within all plants or within all fungi) that is *not* conserved outside of that group (*e.g.* between plants and fungi). All interfaces identified in our analysis are ones that are maintained in the mammalian lineage. Though species-specific interactions are also of great interest, a different approach would be necessary to reveal those regions.

The Gβ gene family was produced by a series of gene duplications. After these events, the new Gβ paralogs likely evolved novel functions as a result of relaxed functional constraint. Eventually, as these new Gβ functions become critical to the organism, these new functional regions become highly conserved regions of the Gβ structure and include the ROIs we identified. We hypothesize that these ROIs represent regions of newly-acquired function in the organism in which they first appeared. For instance, portions of six ROIs presented in Figure 2.4 – ROI2, ROI7, and ROIs10-13 – contain residues that likely bind interactors as previously identified in mutational or

structural studies. The remaining ROIs in Figure 2.4 are in regions that lie outside of known regions presented thus far. Although interactions with many proteins have been determined by mutational studies, many studies focus on known interfaces and therefore indirectly introduce bias. For example, since the Gα binding area is utilized in many Gβ-protein interactions, studies historically focused on these residues. Our *de novo* analysis eliminates this bias, generating a set of potential binding interfaces (ROIs) that are yet to be explored by mutational studies. As with the residues implicated in the PLC-β2 binding interface, these ROIs would make ideal targets for site-directed mutagenesis. In this manner, we can identify new Gβ-interacting proteins as well as to further characterize interactions with existing proteins. Ultimately, this will allow us to characterize new Gβ-mediated signaling pathways.

Acknowledgements

Work in A.M.J.'s lab on the *Arabidopsis* G proteins is supported by the NIGMS (GM065989-01), the DOE (DE-FG02-05ER15671), and the NSF (MCB-0718202 and MCB-0723515). Work in J.S.'s lab is supported by the NIH (5-50921-2311).

Materials and Methods

Ancestral Reconstruction.

To reconstruct plant and fungal ancestors, sequence databases were queried with known plant and mammalian Gβ sequences for full-length plant and fungal Gβ homologs. Additionally, Gβ sequences were also retrieved for two outgroups: entamoeba (*Entamoeba dispar)* and diatom (*Thalassiosira pseudonana*). MSAs were generated in ClustalX (Thompson et al., 1997). Expasy and Genbank IDs for each gene used in the alignment are denoted in Figs. S1 and S2. For each MSA generated, all gaps except those found exclusively in the two outgroups were removed. From the resulting NEXUS file, ancestors were generated using MrBayes (Huelsenbeck et al., 2001; Ronquist and Huelsenbeck, 2003) using the fixed equalin model, using the inverse gamma rate, and sampling 100,000 generations at a frequency of 100. In order to eliminate ancestral residue values chosen with low confidence, only those residues predicted >90% of the time with a maximum value >0.8 were accepted and included in the ancestral sequence. All other residues were assigned a value of "X", as they were too variable to be called with confidence.

Sequence Collection, Alignment, and Phylogeny Generation.

All Gβ sequences from *Drosophila melanogaster*, *Caenorhabditis elegans*, and human (Expasy and Genbank sequence IDs are denoted in Figure 2.S3) were combined with the plant and fungal ancestral sequences and the diatom and entamoeba outgroups to create a MSA (see previous section for alignment and NEXUS file generation). The alignments were made robust by structural comparisons between bovine Gβ1 and mouse Gβ5. MrBayes (Huelsenbeck et al., 2001; Ronquist and Huelsenbeck, 2003) was run

using a fixed equalin model, using the inverse gamma rate, and sampling 1,000,000 generations at a frequency of 100 for 3 independent runs with a burn in of 250,000 generations to generate a consensus phylogenetic tree.

Interface Determination.

In order to identify binding interfaces from structures of Gβ in complex with interacting proteins, Naccess (Hubbard and Thornton, 1993) was used to calculate relative solvent accessibility (RSA) of all Gβ residue side chains both as a monomer and in complex with other proteins. All residue side chains with an RSA decrease > 5% (RSA of Gβ monomer – RSA of complex > 5) between monomer and complex were denoted as being present in the interface. This group of residues also included all Gβ residues that formed hydrogen bonds with the interacting protein as determined in the PyMol Molecular Modeling System (DeLano, 2002). The 5% ΔRSA value was chosen because it is more stringent than the commonly utilized $1Å^2$ change in absolute solvent accessibility (Jones and Thornton, 1996; Jones et al., 2000) but still included all residues which form hydrogen bonds. Additionally, a residue with >=5% RSA is often designated as a surface residue while a residue with <5% RSA is designated as an interior residue (Jones and Thornton, 1996; Jones et al., 2000).

Yeast-Three Hybrid Protein Interaction.

Arabidopsis thaliana Gβ (AGB1) and Gγ1 (AGG1) were cloned into the pBridge vector (Clontech, Palo Alto, CA). *A. thaliana* phosducin (At5g14240) was cloned into the p-ENTR/D-TOPO vector (Invitrogen, Carlsbad, CA) and then recombined into the pACTGW-attR Gateway vector (Nakayama et al., 2002) which contains an activation domain and is compatible with the pBridge vector. The prey (phosducin) was

transformed into yeast strain AH109 which had previously been transformed with the bait (AGB1/AGG1). Both strains (that containing the bait alone and that containing both bait and prey) were grown on nutritionally selective media. Presence of the bait and prey were confirmed by the expression of nutritional markers (positive growth on media lacking tryptophan and leucine, respectively). Interaction was confirmed by the expression of an additional nutritional marker (positive growth on media lacking histidine).

Transfection of COS-7 Cells with Gβ1, Gγ2, and PLC-β2.

The accumulation of [^{3}H]inositol phosphates was measured in transiently transfected COS-7 cells as previously described (Wing et al., 2003). Briefly, COS-7 cells were plated in 12-well dishes at a cell density of 60,000 cells per well in Dulbecco's Modified Eagles Medium (DMEM) supplemented with 10% fetal bovine serum, 10 units / ml penicillin, and 10 units / ml streptomycin. Following incubation for 24 hrs at 37°C in an atmosphere of 95% air / 5% CO_2, cells were transfected with 200 or 300 ng each of the indicated human Gβ1 and Gγ2 DNA in the presence and absence of 30 ng of PLC-β2 and empty vector for a total of 700 ng of DNA per well. DNA was complexed with FuGENE 6 (Roche Applied Sciences, Indianapolis, IN) per manufacturer's protocol prior to transfection. Twenty-four hours post-transfection, the medium was aspirated and cells were metabolically labeled with 1 μCi / well of myo-[2-^{3}H(N)]inositol (American Radiolabeled Chemicals, St. Louis, MO) in inositol-free DMEM for 12 - 16 hours. Subsequently, 10 mM LiCl was added. One hour after incubation with LiCl, reactions were stopped by aspiration of the medium and addition of 50 mM formic acid. Samples were neutralized by the addition of 150 mM NH_4OH, and [^{3}H]inositol phosphates were quantified by Dowex chromatography. Statistical significance for comparisons between

the activation of PLC-β2 by wt Gβ1γ2 and each mutant Gβγ was determined by performing a Student's t-test assuming equal variance between the log-transformed number of [^{3}H]inositol phosphates for each construct. Western blotting was performed to confirm equal expression of each construct in COS-7 cells using an antibody directed toward the c-Myc epitope (Invitrogen) on hGβ1, the HA epitope (Roche Applied Sciences) on Gγ2, and a monoclonal antibody (Santa Cruz, Santa Cruz, CA) against PLC-β2.

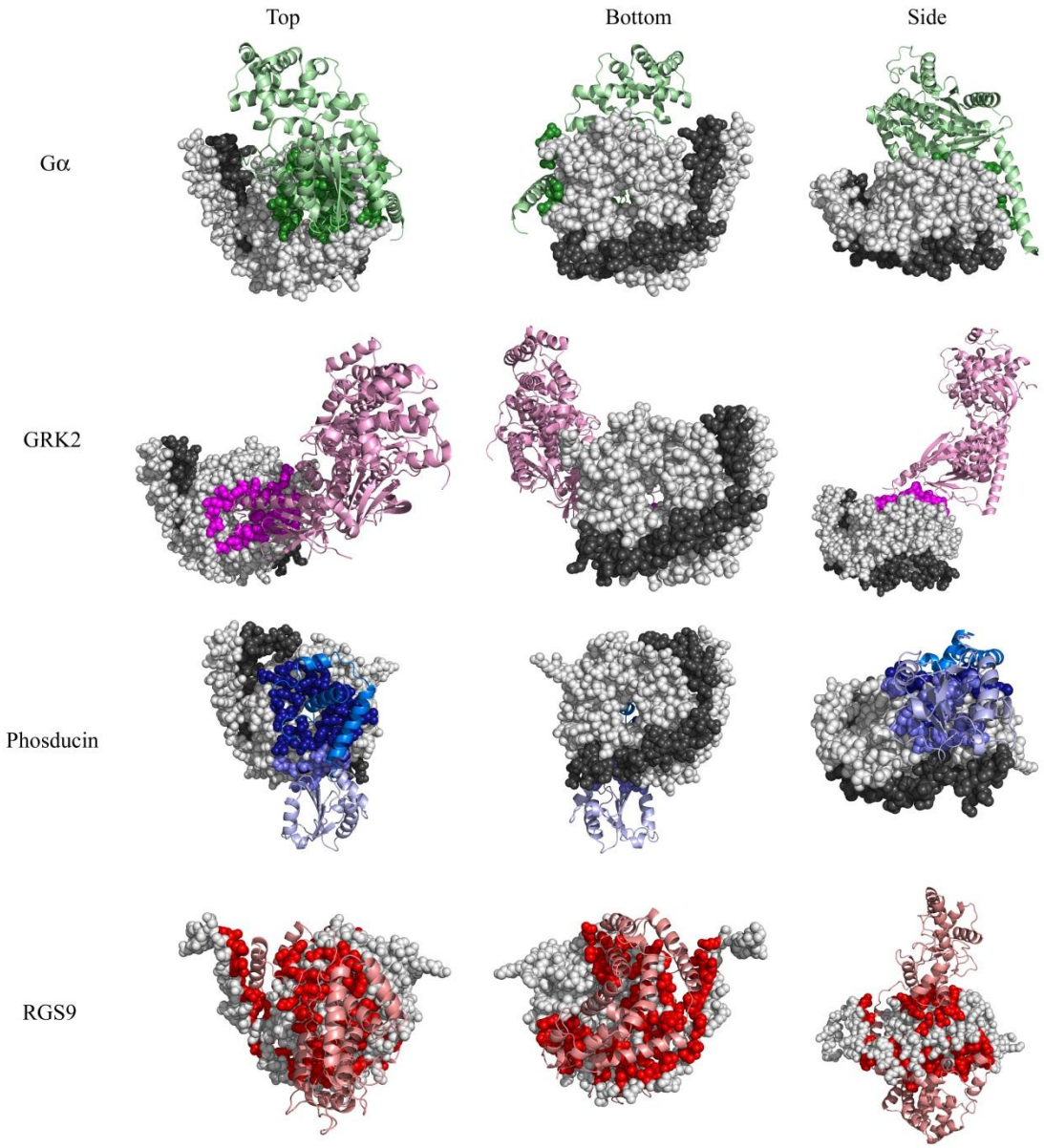


Figure 2.1. Binding interfaces of four Gβ-interacting proteins as determined by crystal structures. Top, bottom, and side views of Gβ or the Gβγ dimer (Gβ in light gray spheres, Gγ in dark gray spheres) bound to four different interacting proteins (ribbon): Gβ1-Gα in green (1got.pdb), Gβ1-GRK2 in magenta (1omw.pdb), Gβ1-phosducin in blue (N-terminal domain is dark, C-terminal domain is light, 1a0r.pdb), and Gβ5-RGS9 in red (2pbi.pdb). Gβ residues that contact each interacting protein are colored accordingly. Binding contacts were determined by evaluating the solvent accessibility difference between single molecules and those in complex.

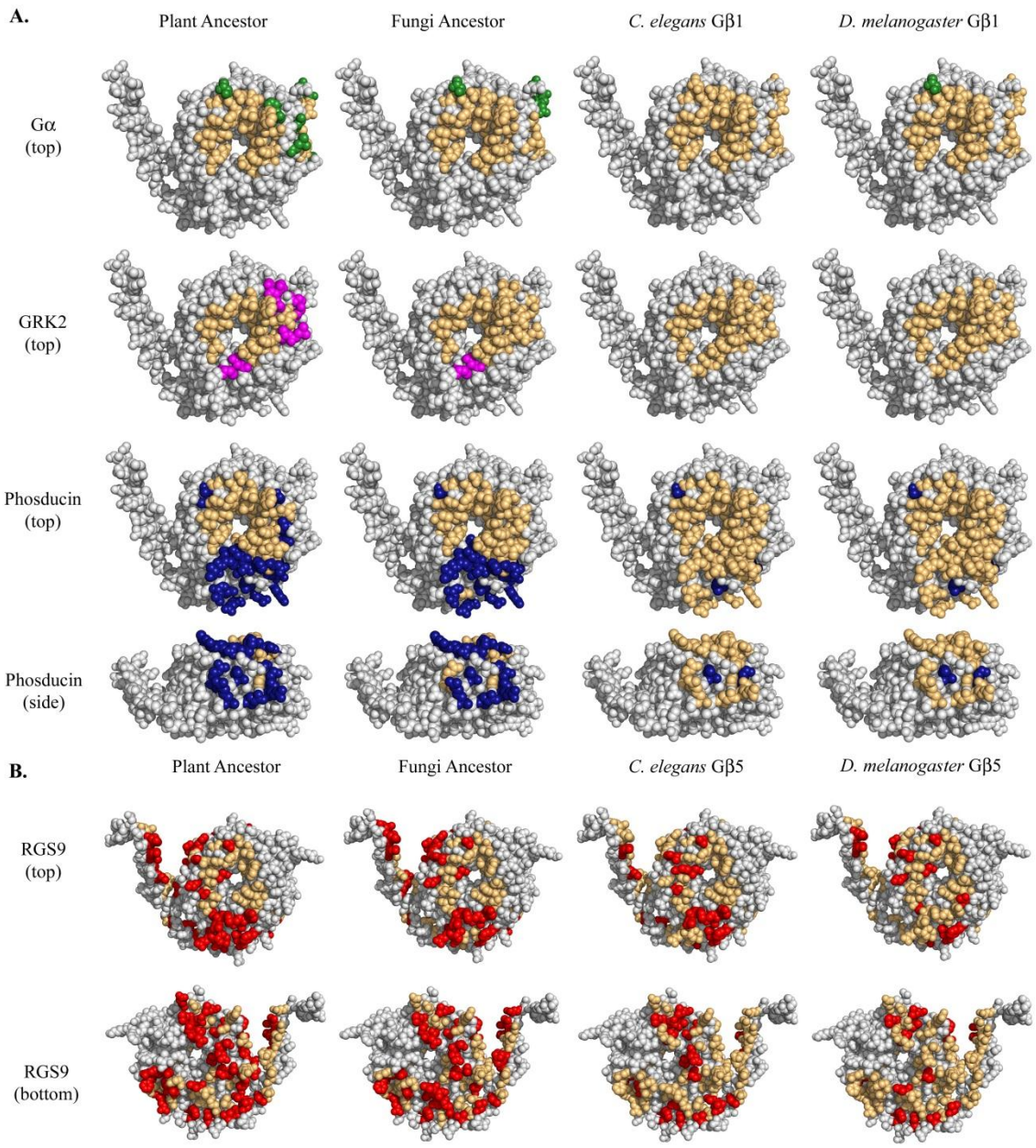


Figure 2.2. Conservation within known binding interfaces based on bovine Gβ1 crystal structures. Conservation of the (A) Gα, GRK2, and phosducin binding interfaces on Gβ1 (1got.pdb) and the (B) RGS9 binding interface on Gβ5 (2pbi.pdb) in four reference organisms. While the interface is comprised of all colored residues, conserved residues are colored light orange while non-conserved residues are colored green (Gα), blue (phosducin), magenta (GRK2), or red (RGS9). Only residues in the binding area were analyzed.

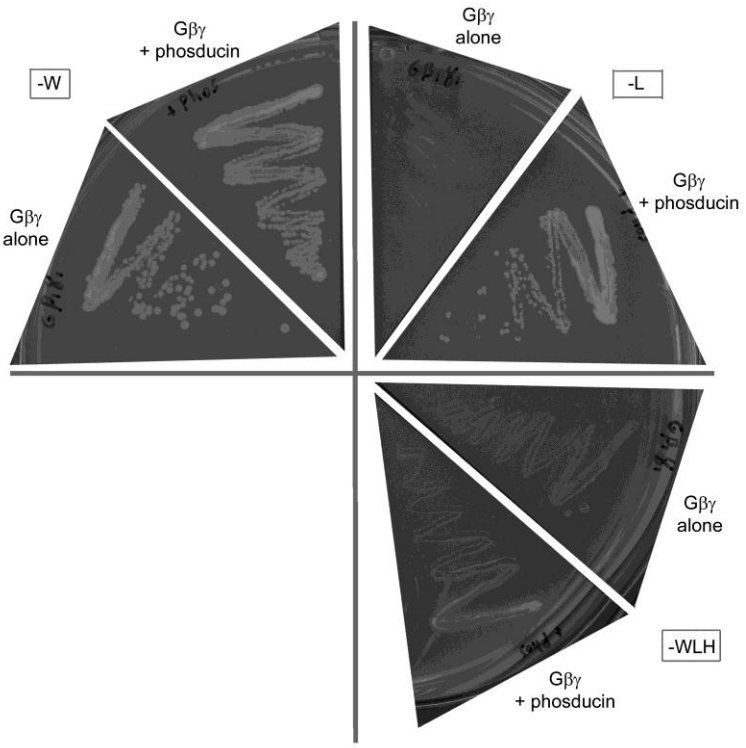


Figure 2.3. An *Arabidopsis* phosducin and Gβ interact physically. Growth of yeast strain AH109 containing the genes indicated (AtGβ1γ1 alone or AtGβ1γ1 and phosducin) on yeast dropout media. Media missing tryptophan (-W) selects for the AtGβ1γ1 vector, resulting in positive growth for both genotypes. Media missing leucine (-L) selects for the phosducin vector, resulting in no growth for the strain lacking phosducin and positive growth for the strain containing phosducin. Media missing leucine, tryptophan, and histidine (-LWH) selects for a positive interaction between the two genes, resulting in no growth for the strain containing AtGβ1γ1 alone and positive growth for the strain containing both AtGβ1γ1 and phosducin. The latter growth indicates that the two genes physically interact.

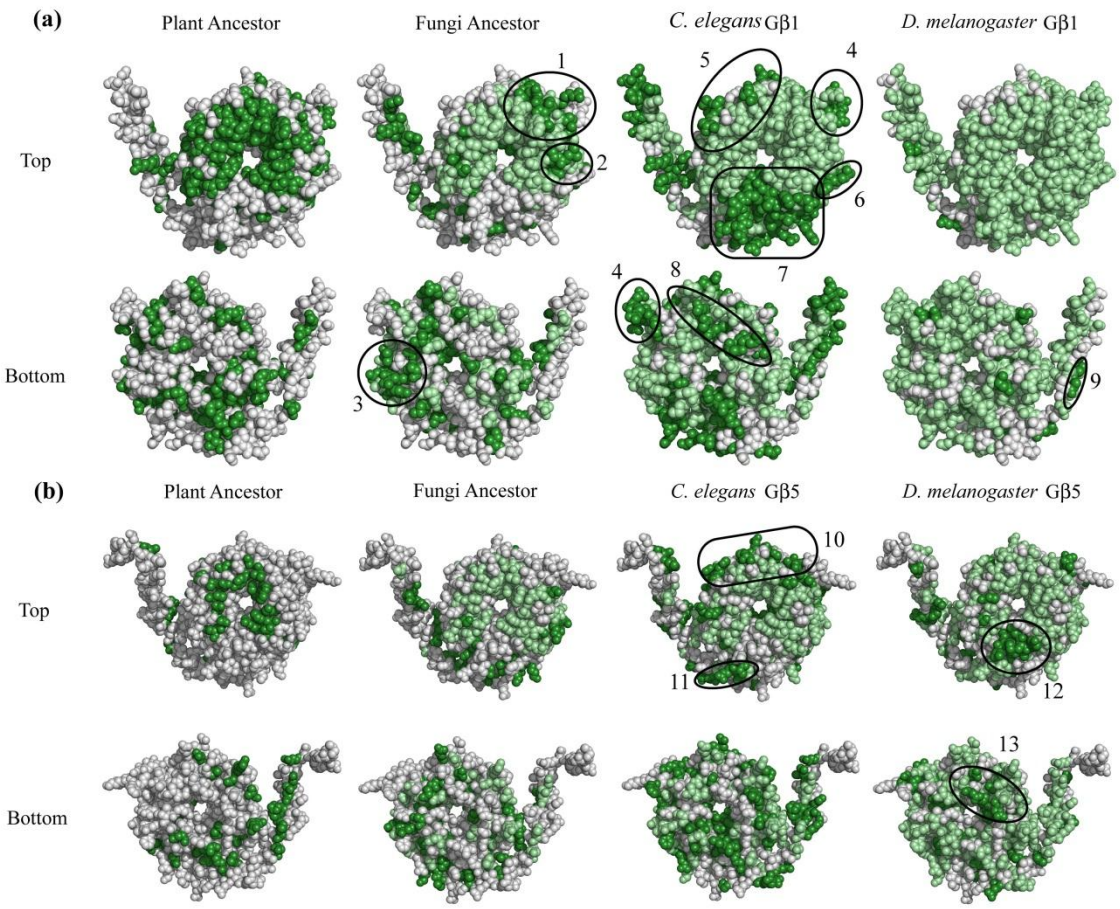


Figure 2.4. Gβ regions of interest as determined by a comparative evolutionary analysis. Three views (top, bottom, side) of the evolution of conserved regions in the plant ancestor, fungal ancestor, *C. elegans* Gβ, and *D. melanogaster* Gβ (a, Gβ1 [1got.pdb] and b, Gβ5 [2pbi.pdb]). The dark green regions of the plant ancestor highlight primordial function and form region of interest (ROI) 0. In all other organisms, newly conserved residues (those that matched the mammalian value) were colored dark green, while conserved residues present in a previous organism were colored light green. Thus, dark green patches represent acquisition of new function. ROIs are indicated and corresponding residues are listed in Table 2.1.

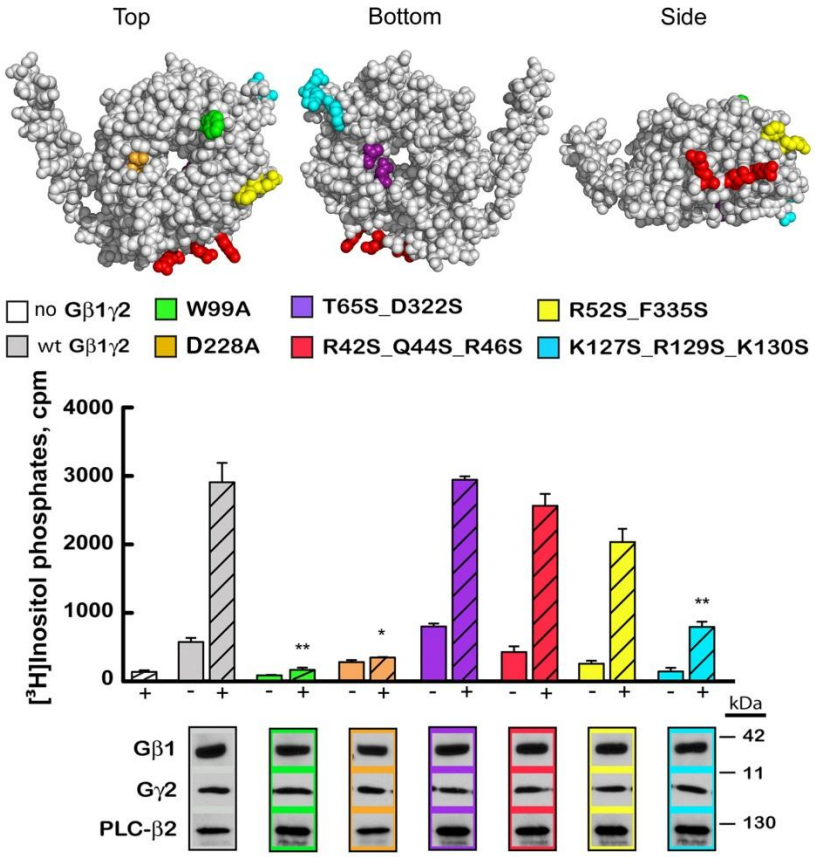


Figure 2.5. Regions of Gβ involved in PLC-β2 activation. Groups of residues mutated on the Gβ surface [1got.pdb] are colored and coordinate with colors in the bar graph and associated blots. Wild-type and mutant Gβ1 subunits were tested in the absence (-) or presence (+) of PLC-β2 for their ability to activate PLC-β2 (*p<0.01 **p<0.005, error bars represent standard error). Immunoblot analysis confirmed equal expression of all proteins utilized.

Table 2.1. ROI residue numbers. Mammalian residue positions for each Gβ residue within the regions of interest highlighted in Figure 2.4.

ROI#	Residues	First conserved in
1	R96, S97, S98, I120, R134, E138, E172	Fungi Gβ1
2	L55, A56, S334	Fungi Gβ1
3	D66, R68, Y85, N88, V90, Y105	Fungi Gβ1
4	K127, R129, E130, N132	*C. elegans* Gβ1
5	T143, T173, Q175, Q176, T181, T184, M217	*C. elegans* Gβ1
6	R46, T47, R52, D312, F335	*C. elegans* Gβ1
7	R42, Q44, D267, N268, I269, I270, C271, G272, I273, D290, D291, N293, N295, V307, A309	*C. elegans* Gβ1
8	L152, D153, N155, D195, R197, L210	*C. elegans* Gβ1
9	D20, A24	*D. melanogaster* Gβ1
10	T102, P104, T106, N141, K146, N154	*C. elegans* Gβ5
11	E43, K279, E280, S281	*C. elegans* Gβ5
12	F284, N303, Y305	*D. melanogaster* Gβ5
13	N163, L203, P205, E207	*D. melanogaster* Gβ5

Supplementary Materials

The following figures were published as supplementary materials and can be viewed here or on the JMB website. Two of the protein sequence alignments were used to generate the fungal and plant ancestors to Gβ, and the third alignment was used to identify conserved residues between the groups of representative Gβ sequences. The phylogenetic tree displays the evolutionary relationships between the representative Gβ sequences.

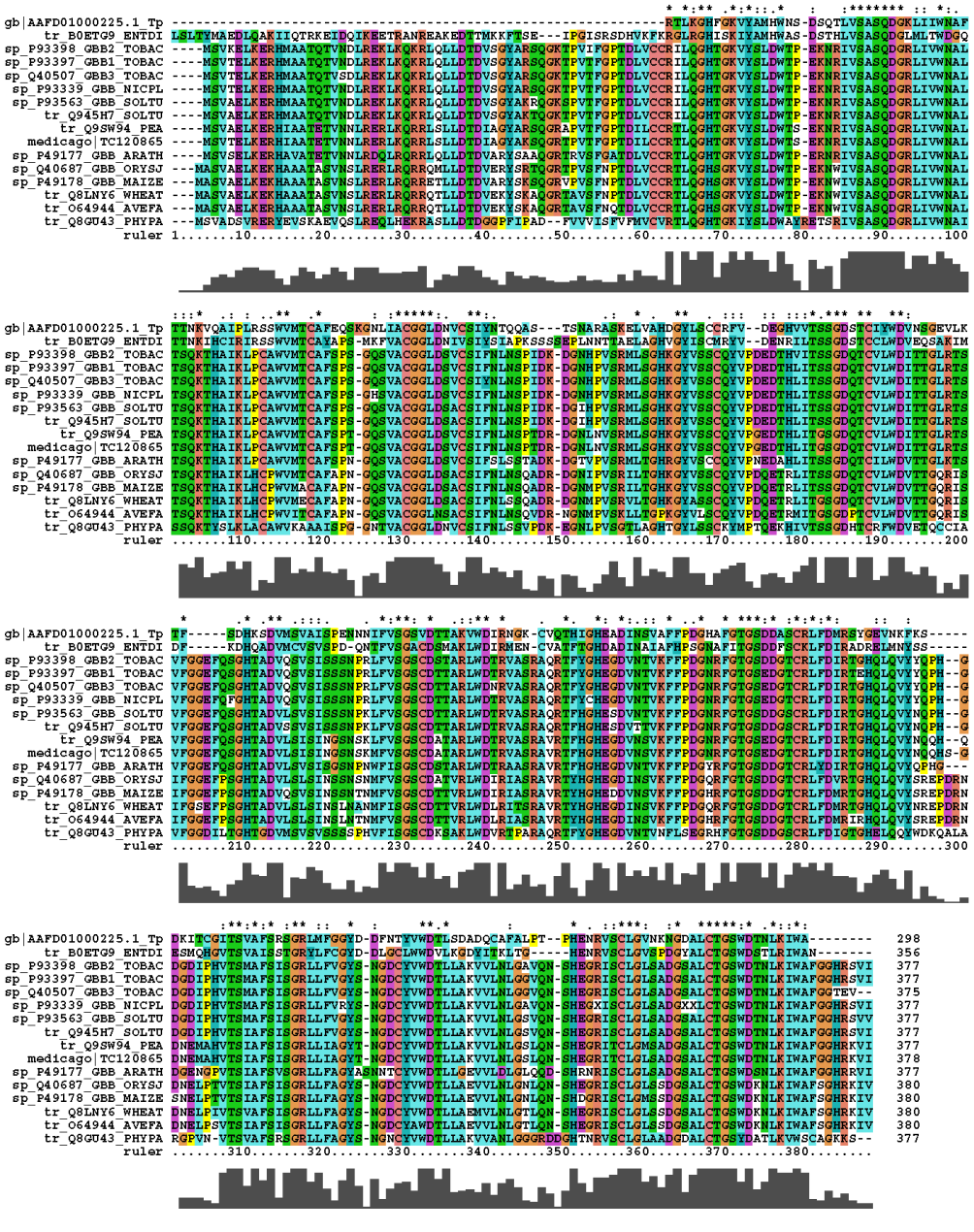
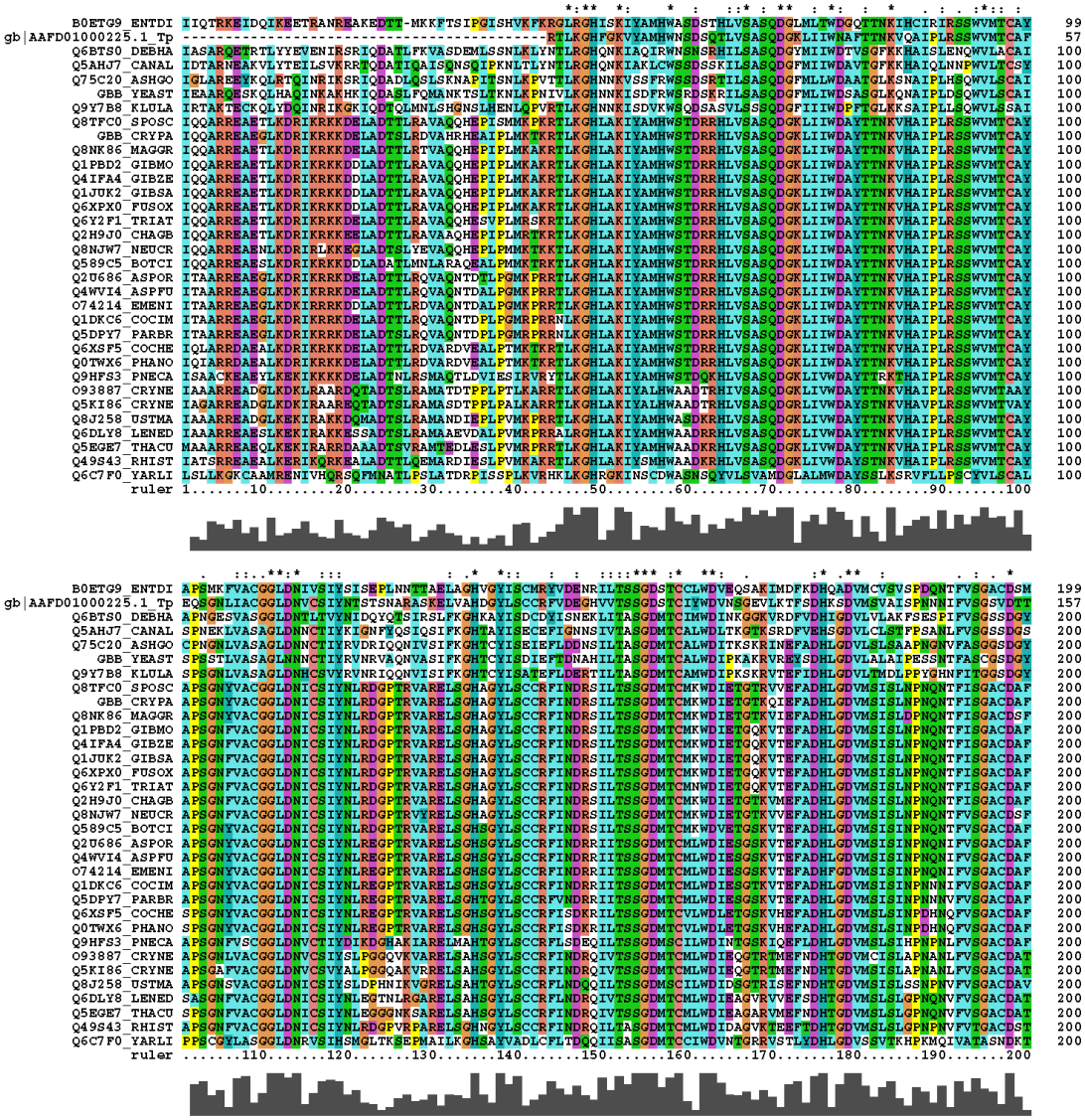


Figure 2.S1. Alignment used to generate the plant Gβ ancestor. Plant sequences were aligned using Clustal X. Diatom and entamoeba Gβ sequences served as outgroups during ancestral reconstruction. Expasy and Genbank accession numbers are indicated within the sequence name in the alignment. Residues are colored according to their identity and symbols above each residue indicate residues that are strongly conserved ("*" full conservation of a single residue, ":" full conservation within a "strong" group, "." full conservation within a "weak" group). The bar graph beneath the alignment represents the conservation score at each position in the alignment.



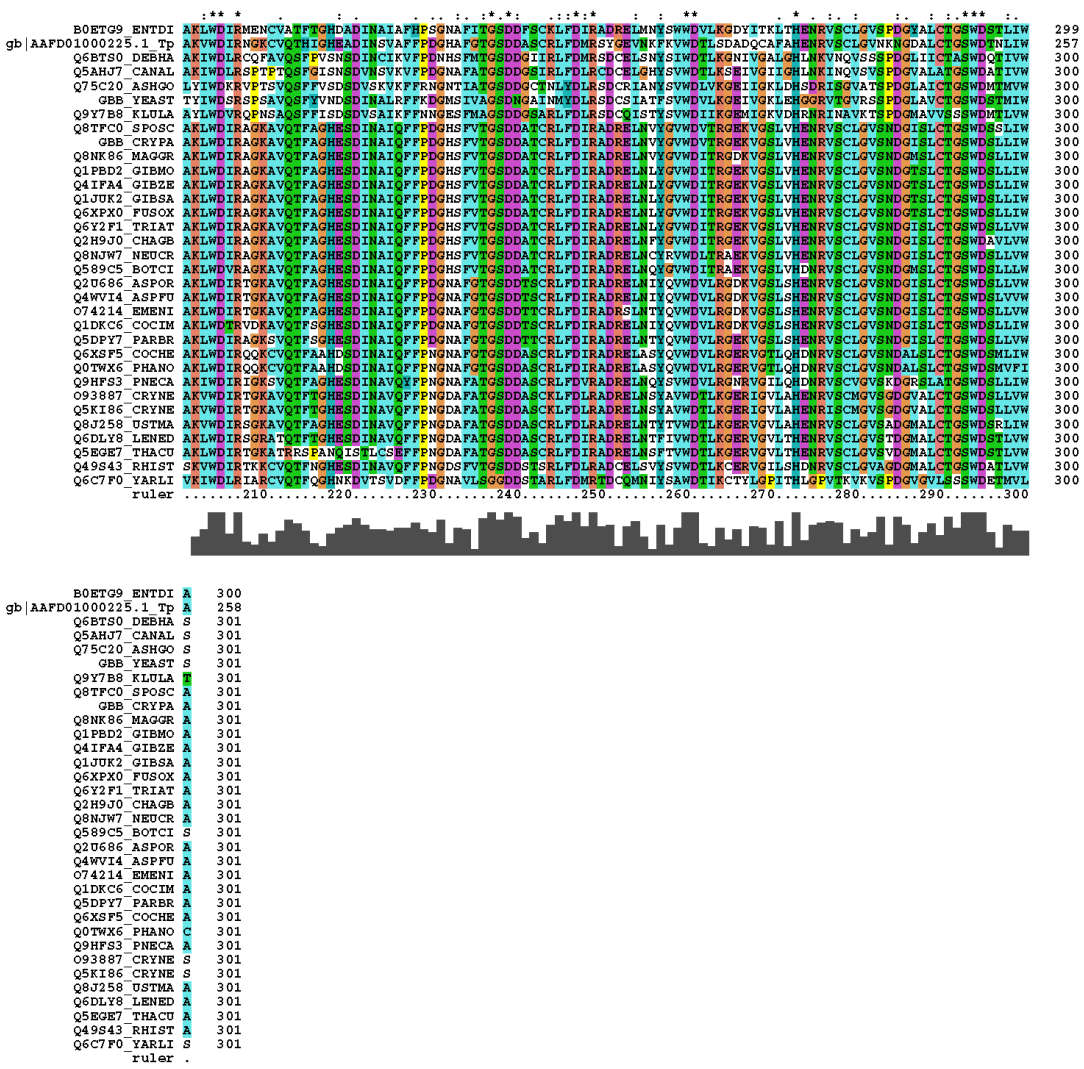


Figure 2.S2. Alignment used to generate the fungal Gβ ancestor. Fungal sequences were aligned using Clustal X. Diatom and entamoeba Gβ sequences served as outgroups during ancestral reconstruction. Expasy and Genbank accession numbers are indicated within the sequence name in the alignment. Residues are colored according to their identity and symbols above each residue indicate residues that are strongly conserved ("*" full conservation of a single residue, ":" full conservation within a "strong" group, "." full conservation within a "weak" group). The bar graph beneath the alignment represents the conservation score at each position in the alignment.

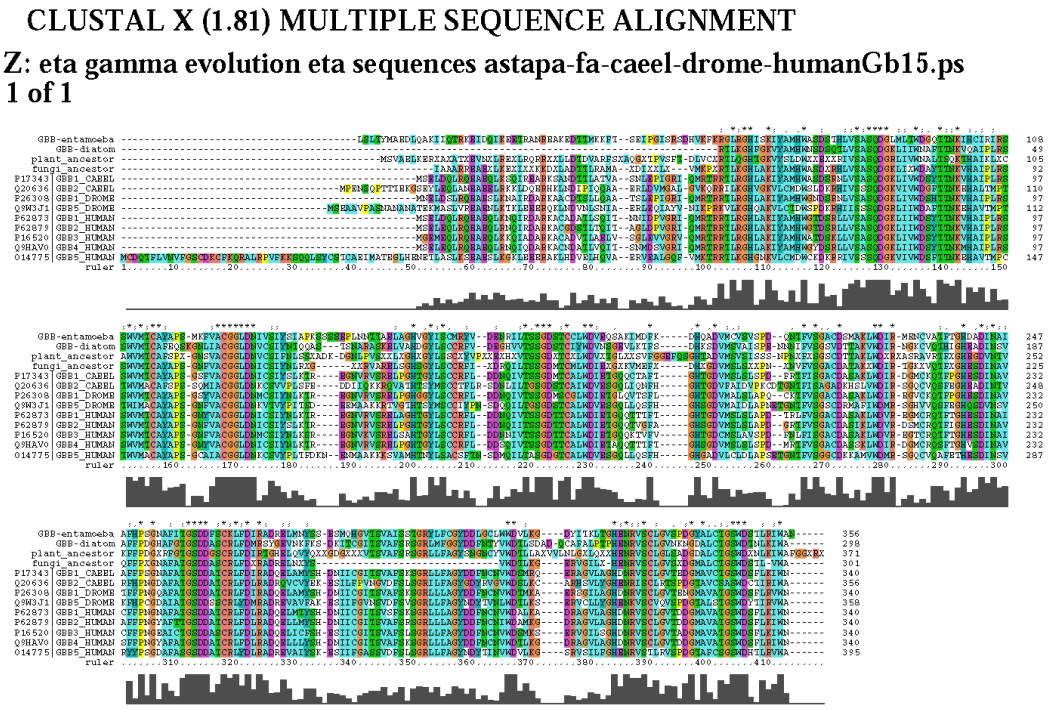


Figure 2.S3. Alignment of representative Gβ sequences. Gβ sequences from the plant ancestor, fungal ancestor, *C. elegans*, *D. melanogaster*, and human were aligned in ClustalX. These alignments were corrected to match the structural alignment between mammalian Gβ1 and Gβ5. Expasy and Genbank accession numbers are indicated within the sequence name in the alignment. Residues are colored according to their identity and symbols above each residue indicate residues that are strongly conserved ("*" full conservation of a single residue, ":" full conservation within a "strong" group, "." full conservation within a "weak" group). The bar graph beneath the alignment represents the conservation score at each position in the alignment.

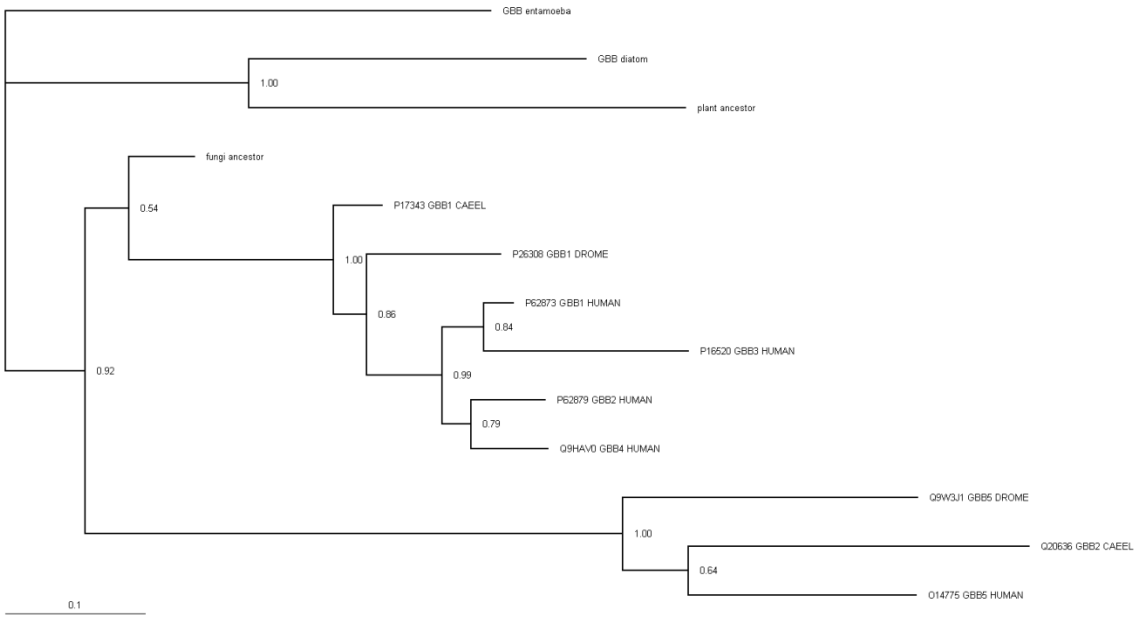


Figure 2.S4. Phylogeny of representative Gβ sequences. The MSA (Figure 2.S3) was used to generate a consensus phylogenetic tree using MrBayes with *Thalassioria* (diatom) and *Entamoeba* serving as outgroups. Metazoan Gβ sequences form two main groups, Gβ1-like and Gβ5-like. Confidence intervals are noted on the phylogeny.

References

Armon, A., Graur, D., and Ben-Tal, N. (2001). ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. J Mol Biol **307,** 447-463.

Blaauw, M., Knol, J.C., Kortholt, A., Roelofs, J., Ruchira, Postma, M., Visser, A.J., and van Haastert, P.J. (2003). Phosducin-like proteins in Dictyostelium discoideum: implications for the phosducin family of proteins. Embo J **22,** 5047-5057.

Bonacci, T.M., Ghosh, M., Malik, S., and Smrcka, A.V. (2005). Regulatory interactions between the amino terminus of G-protein betagamma subunits and the catalytic domain of phospholipase Cbeta2. J Biol Chem **280,** 10174-10181.

Cabrera-Vera, T.M., Vanhauwe, J., Thomas, T.O., Medkova, M., Preininger, A., Mazzoni, M.R., and Hamm, H.E. (2003). Insights into G protein structure, function, and regulation. Endocr Rev **24,** 765-781.

Cheever, M.L., Snyder, J.T., Gershburg, S., Siderovski, D.P., Harden, T.K., and Sondek, J. (2008). Crystal structure of the multifunctional Gbeta5-RGS9 complex. Nat Struct Mol Biol **15,** 155-162.

Davis, T.L., Bonacci, T.M., Sprang, S.R., and Smrcka, A.V. (2005). Structural and molecular characterization of a preferred protein interaction surface on G protein beta gamma subunits. Biochemistry **44,** 10593-10604.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.

Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.S., Iyengar, R., Miller, R.J., Jan, L.Y., Lefkowitz, R.J., and Hamm, H.E. (1998). Molecular basis for interactions of G protein betagamma subunits with effectors. Science **280,** 1271-1274.

Fukuto, H.S., Ferkey, D.M., Apicella, A.J., Lans, H., Sharmeen, T., Chen, W., Lefkowitz, R.J., Jansen, G., Schafer, W.R., and Hart, A.C. (2004). G protein-coupled receptor kinase function is essential for chemosensation in C. elegans. Neuron **42,** 581-593.

Gaudet, R., Bohm, A., and Sigler, P.B. (1996). Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin. Cell **87,** 577-588.

Gu, X., and Vander Velden, K. (2002). DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. Bioinformatics **18,** 500-501.

Hubbard, S.J., and Thornton, J.M. (1993). 'NACCESS,' computer program. Department of Biochemistry and Molecular Biology, University College London.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., and Bollback, J.P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. Science **294,** 2310-2314.

Johnston, C.A., Kimple, A.J., Giguere, P.M., and Siderovski, D.P. (2008). Structure of the parathyroid hormone receptor C terminus bound to the G-protein dimer Gbeta1gamma2. Structure **16,** 1086-1094.

Jones, A.M., and Assmann, S.M. (2004). Plants: the latest model system for G-protein research. EMBO Rep **5,** 572-578.

Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. Proc Natl Acad Sci U S A **93,** 13-20.

Jones, S., Marin, A., and Thornton, J.M. (2000). Protein domain interfaces: characterization and comparison with oligomeric protein interfaces. Protein Eng **13,** 77-82.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E., and Sigler, P.B. (1996). The 2.0 A crystal structure of a heterotrimeric G protein. Nature **379,** 311-319.

Li, Y., Sternweis, P.M., Charnecki, S., Smith, T.F., Gilman, A.G., Neer, E.J., and Kozasa, T. (1998). Sites for Galpha binding on the G protein beta subunit overlap with sites for regulation of phospholipase Cbeta and adenylyl cyclase. J Biol Chem **273,** 16265-16272.

Lichtarge, O., Bourne, H.R., and Cohen, F.E. (1996a). An evolutionary trace method defines binding surfaces common to protein families. J Mol Biol **257,** 342-358.

Lichtarge, O., Bourne, H.R., and Cohen, F.E. (1996b). Evolutionarily conserved Galphabetagamma binding surfaces support a model of the G protein-receptor complex. Proc Natl Acad Sci U S A **93,** 7507-7511.

Lodowski, D.T., Pitcher, J.A., Capel, W.D., Lefkowitz, R.J., and Tesmer, J.J. (2003). Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma. Science **300,** 1256-1262.

Loew, A., Ho, Y.K., Blundell, T., and Bax, B. (1998). Phosducin induces a structural change in transducin beta gamma. Structure **6,** 1007-1019.

Madabushi, S., Yao, H., Marsh, M., Kristensen, D.M., Philippi, A., Sowa, M.E., and Lichtarge, O. (2002). Structural clusters of evolutionary trace residues are statistically significant and common in proteins. J Mol Biol **316,** 139-154.

McCudden, C.R., Hains, M.D., Kimple, R.J., Siderovski, D.P., and Willard, F.S. (2005). G-protein signaling: back to the future. Cell Mol Life Sci **62,** 551-577.

Miller, K.G., Emerson, M.D., and Rand, J.B. (1999). Goalpha and diacylglycerol kinase negatively regulate the Gqalpha pathway in C. elegans. Neuron **24,** 323-333.

Milligan, G., and Kostenis, E. (2006). Heterotrimeric G-proteins: a short history. Br J Pharmacol **147 Suppl 1,** S46-55.

Myung, C.S., Lim, W.K., DeFilippo, J.M., Yasuda, H., Neubig, R.R., and Garrison, J.C. (2006). Regions in the G protein gamma subunit important for interaction with receptors and effectors. Mol Pharmacol **69,** 877-887.

Nakayama, M., Kikuno, R., and Ohara, O. (2002). Protein-protein interactions between large proteins: two-hybrid screening using a functionally classified library composed of long cDNAs. Genome Res **12,** 1773-1784.

Ohno, S. (1970). Evolution by gene duplication. (Berlin: Springer).

Panchenko, M.P., Saxena, K., Li, Y., Charnecki, S., Sternweis, P.M., Smith, T.F., Gilman, A.G., Kozasa, T., and Neer, E.J. (1998). Sites important for PLCbeta2 activation by the G protein betagamma subunit map to the sides of the beta propeller structure. J Biol Chem **273,** 28298-28304.

Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19,** 1572-1574.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25,** 4876-4882.

van der Linden, A.M., Simmer, F., Cuppen, E., and Plasterk, R.H. (2001). The G-protein beta-subunit GPB-2 in Caenorhabditis elegans regulates the G(o)alpha-G(q)alpha signaling network through interactions with the regulator of G-protein signaling proteins EGL-10 and EAT-16. Genetics **158,** 221-235.

Wing, M.R., Snyder, J.T., Sondek, J., and Harden, T.K. (2003). Direct activation of phospholipase C-epsilon by Rho. J Biol Chem **278,** 41253-41258.

Yoshikawa, D.M., Bresciano, K., Hatwar, M., and Smrcka, A.V. (2001). Characterization of a phospholipase C beta 2-binding site near the amino-terminal coiled-coil of G protein beta gamma subunits. J Biol Chem **276,** 11246-11251.

Zheng, Y., Xu, D., and Gu, X. (2007). Functional divergence after gene duplication and sequence-structure relationship: a case study of G-protein alpha subunits. J Exp Zoolog B Mol Dev Evol **308,** 85-96.

CHAPTER 3

Aci-Reductone Dioxygenase 1 (ARD1) is an Effector of the Heterotrimeric G Protein Beta Subunit in *Arabidopsis*

Preface

For the third chapter, I have included my first-author paper that was submitted to the Journal of Biological Chemistry and entitled "Aci-reductone dioxygenase 1 (ARD1) is an effector of the heterotrimeric G protein beta subunit in *Arabidopsis.*" The co-authors on this paper were Helen X. Wang, Iva Perovic, Aditi Deshpande, Thomas C. Pochapsky, Brenda R. S. Temple, Stephanie N. Hicks, T. Kendall Harden, and Alan M. Jones. Brenda Temple assisted with the structural and evolutionary analyses and the modeling of ARD1. Iva Perovic and Aditi Deshpande performed the ARD1 enzymatic assays under the direction of Tom Pochapsky. Helen Wang performed the initial enhancer screen and identified and cloned ARD1. Stephanie Hicks and Ken Harden assisted in the production of viruses and in the maintenance of the Sf9 cells used to generate the AGB1 wildtype and mutant proteins. This work was performed in the lab and under the direction of Alan Jones.

Abstract

Components of the heterotrimeric G protein complex are conserved from plants to mammals, but the complexity of each system varies. The plant *Arabidopsis thaliana* contains one Gα, one Gβ, and at least two Gγ subunits, allowing it to form two versions of the heterotrimer. This simple plant model is ideal for genetic studies because mammalian systems can potentially contain thousands of unique heterotrimers. The activation of these complexes promotes interactions between both the Gα subunit and the Gβγ dimer with enzymes and scaffolds to propagate signaling to the cytoplasm. However, while effectors of Gα and Gβ are known in mammals, no Gβ effectors were previously known in plants. Using a genetic suppressor screen, we found that overexpression of aci-reductone dioxygenase 1 (ARD1) suppresses the two-day-old, etiolated phenotype of the Gβ-null mutant *agb1-2*. ARD1 is homologous to prokaryotic and eukaryotic ARD proteins; one function of ARDs is to operate in the methionine salvage pathway. ARD1 is an active metalloenzyme and AGB1 and ARD1 both control embryonic hypocotyl length by modulating cell division; they also may contribute to the production of the hormone ethylene, an indirect product of the methionine salvage pathway. ARD1 physically interacts with AGB1 and ARD enzymatic activity is stimulated by AGB1 *in vitro*; therefore, ARD1 is an effector of AGB1. The binding interface on AGB1 was deduced using a comparative evolutionary approach and tested using recombinant AGB1 mutants. In addition, a possible mechanism for AGB1 activation of ARD1 activity was tested using directed mutations in a loop near the substrate binding site.

Introduction

Arabidopsis plants lacking an AGB1 mRNA transcript (*agb1-2*) have a wide array of developmental phenotypes. Mature *agb1-2* plants display aberrant leaf shape, silique morphology, and increased root mass (Ullah et al., 2003) and are hypersensitive to infection by the necrotrophic pathogens *P. cucumernia* and *F. oxysporum* (Llorente et al., 2005; Trusov et al., 2008). Two-day-old etiolated *agb1-2* seedlings have a shorter, thicker hypocotyl and a more open apical hook than wildtype seedlings (Ullah et al., 2001; Ullah et al., 2003). The shortened hypocotyl in *agb1*-2 seedlings is due to fewer cells (Ullah et al., 2001; Ullah et al., 2003), and therefore a decrease in cell division. The apical hook morphology is controlled by a number of factors including the hormone ethylene. Increased ethylene production is one factor that promotes the closure of the apical hook (Woeste and Kieber, 1998).

Although we previously identified proteins that interact genetically (SGB1 (Wang et al., 2006)) and physically (NDL1 (Mudgil et al., 2009)) with AGB1, to date no AGB1 effectors (proteins whose activity are regulated by AGB1) have been identified in plants. This is in stark contrast to the wide array of studies regarding mammalian Gβ signaling. For example, in mammals, many Gβ effectors have been identified, including adenylyl cyclase 2 (AC2), phospholipase C β2 (PLC-β2), and calcium and potassium channels (Smrcka, 2008). However, no protein that interacts with AGB1 in plants has been shown to play a direct role in the modulation of cell division. In order to identify potential effectors in the AGB1 signaling pathway, we utilized an activation-tagging approach to randomly create dominant suppressors of *agb1-2* (Weigel et al., 2000). Specifically, we

screened genes whose upregulation would restore the cell division phenotype displayed in the two-day-old *agb1-2* hypocotyl.

Results

Overexpression of the coding region of At4g14716 (formally SGB3) suppresses the agb1-2 etiolated phenotype.

In order to identify genetic interactors of *AGB1*, *AGB1* null (*agb1-2)* seeds were transformed with a 35S enhancer that randomly inserted into the genome. The resulting transgenic population was screened for genes whose overexpression suppressed the shortened hypocotyls and open apical hook present in two-day-old etiolated *agb1-2* seedlings (Figure 3.1A, B) (Weigel et al., 2000; Ullah et al., 2003). Eight genetic loci suppressed this phenotype to varying degrees, and these loci were named Suppressor of G Beta 1-8. The suppressor screen and one of the resulting genetic suppressors, *SGB1*, was previously described (Wang et al., 2006). The current study focuses on *SGB3*, which fully rescued the hypocotyl length phenotype and partially rescued the apical hook opening phenotype. The genomic fragment containing the enhancer was isolated by plasmid rescue, and the genetic position of the enhancer was determined by sequencing. Because the 35S enhancer used can stimulate expression of genes within a 10 kb range (5 kb in each direction of the enhancer insertion position) (Weigel et al., 2000), the transcript levels of the four adjacent genes (within the 10 kb region surrounding the enhancer) were tested to determine which genes had higher expression levels due to the enhancer (Figure 3.1C). The only gene displaying an elevated mRNA transcript was At4g14716 (Arabidopsis Genome Initiative accession). In order to show that increased expression of At4g14716 was responsible for the rescued *agb1-2* phenotype that was

observed in the enhancer population, we ectopically expressed the coding region of At4g14716 under the control of the strong 35S promoter into both *agb1-2* and wildtype backgrounds. As shown in Figure 3.1D, over- and ectopic expression of the At4g14716 coding region rescued the etiolated hook and hypocotyl phenotypes observed in *agb1-2* seedlings. Additionally, the overexpression of the At4g14716 coding region in the presence of *AGB1* (wildtype plants) did not affect the two-day-old, etiolated apical hook opening or hypocotyl length, indicating that the At4g14716 overexpression phenotype requires the loss of *AGB1*. These data show that *AGB1* and At4g14716 interact genetically to control hypocotyl length and apical hook opening.

At4g14716 encodes ARD1, an aci-reductone dioxygenase.

A BLAST analysis revealed that *SGB3* encodes a gene previously annotated as *ARD1*, an aci-reductone dioxygenase-like protein. ARD1 has high similarity to previously identified eukaryotic aci-reductone dioxygenase proteins (Figure 3.S1A). ARD1 is 79% identical and 92% similar to *Oryza sativa* (rice) ARD1 (OsARD1, e value = $4.0x10^{-91}$) and 63% identical and 77% similar to *Mus musculus* (mouse) ARD (MmADI1, e value = $5.0x10^{-65}$). ARD1 is 28% identical and 49% similar to a prokaryotic ARD from the bacterium *Klebsiella oxytosa* (KoARD, e value = $1.0x10^{-7}$). There are two available ARD atomic structures, one of KoARD and one of MmADI1. Based on the higher sequence similarity between ARD1 and MmADI1, we utilized the MmADI1 crystal structure (PDB ID 1VR3) (Xu et al., 2006) as a template and threaded ARD1 onto it to generate a homology model (Figure 3.S1B). The structure was robust (homology score = 0.74) and as shown in Figure 3.S1B, most of the surface and active site residues were conserved between plants and mammals (pink residues). Additionally, the three

histidine residues and one glutamic acid required for metal binding in the ARD1 active site (Chai et al., 2008) were fully conserved (dark blue residues).

The enzymatic activity of ARD proteins was previously characterized in bacteria, plants, and mammals. ARD catalyzes a committed step in the methionine salvage pathway, which recycles methionine from S-adenosylmethionine (SAM) (Pirkov et al., 2008) (Figure 3.1E). This pathway is important in that it provides the organism with a source of methionine under limiting conditions, regulates the production of polyamines, and in plants, allows for the production of ethylene (Lin et al., 2005). ARD proteins catalyze two distinct reactions dependent upon which divalent metal ion is bound in the active site (Dai et al., 1999; Ju et al., 2006). Fe-bound ARD catalyzes the on-pathway reaction which converts aci-reductone to the keto-acid α-keto-γ-methylthiobutyrate (the methionine precursor). Ni-ARD catalyzes an off-pathway reaction whose products are γ-(methylthio)propionate, carbon monoxide, and formate (Pochapsky et al., 2002). The different ARD activities are measured *in vitro* by monitoring the accumulation of CO (off-pathway) or keto-acid (on-pathway) (Wray and Abeles, 1995; Dai et al., 2001; Sauter et al., 2005).

ARD1 function in vivo.

ARD1 T-DNA insertion mutants were obtained to determine the loss-of-function, hypocotyl phenotype. Two of the three *ARD1* alleles, *ard1-1* and *ard1-3* were transcript-null, and *ard1-2* likely created a truncated transcript as determined by RT-PCR (Figure 3.2A). Two-day-old, etiolated seedlings lacking a functional ARD1 transcript mimicked the *agb1-2* phenotype in that they had short hypocotyls, and they showed an apical hook opening that was intermediate to *agb1-2* and wildtype plants (Figure 3.2B). We compared

the estimated number of epidermal cells in the two-day-old etiolated seedlings and found that *ard1-1,* like *agb1-2,* contains fewer cells (~10 and ~9 cells, respectively) than wildtype plants (~20 cells).

These results prompted the hypothesis that AGB1 regulates ARD1 activity to control cell division. Regulation could occur at different levels, for example, by control of ARD1 subcellular location or by direct control of ARD1 catalytic properties. To test the former, the localization of ARD1-GFP was determined in both the wildtype and *agb1*-2 backgrounds, and no major difference in patterns was detected using transient co-expression in protoplasts (Figure 3.S2). Therefore, AGB1 does not affect the cytoplasmic localization of ARD1.

One of the products of the methionine salvage pathway in plants is the gaseous hormone ethylene (Figure 3.1E). When etiolated seedlings are grown in the presence of ethylene, they display a set of phenotypes called the triple response, shorter hypocotyls, exaggerated closure of the apical hook, and radially-expanded hypocotyls (Woeste and Kieber, 1998). The steady-state level of ethylene in *agb1* and *ard1* mutant seedlings was determined because 1) ARD1 may operate in the methionine salvage pathway based on its homology with known ARD proteins (Figure 3.S1), 2) *AGB1* positively regulates *ARD1* genetically, and 3) two-day-old etiolated *agb1-2* and *ard1* mutants display some aspects of the triple response. *ard1* null mutants display one triple response trait (short hypocotyls), leading to the prediction that ethylene levels in *ard1* mutants may be elevated. On the other hand, another *ard1* triple response trait (open hooks) led to the prediction that ethylene levels are lower than wildtype seedlings. The steady-state level of ethylene in two-day-old etiolated seedlings was slightly reduced in *agb1-2*, *ard1-1*,

ard1-2, and *ard1-3* compared to wildtype (Figure 3.2C). These data correlate with the two-day-old apical hook phenotypes of these plants (Figure 3.2D).

ARD1 physically interacts with AGB1.

We showed that *ARD1* genetically interacts with *AGB1* because *ARD1* overexpression rescues the *agb1-2* etiolation phenotype (Figure 3.1). To test for physical interaction, yeast-three-hybrid (Y3H) genetic complementation was performed. To confirm that each construct could be expressed in yeast cells, single transformants were grown on an appropriate nutritional dropout medium (Δleucine for ARD1, Δtryptophan for AGB1) (Figure 3.3A). ARD1 was expressed in AH109 yeast cells in the presence or absence of the AGB1-AGG1 heterodimer. When grown on nutritional media selecting for a positive interaction between the three proteins (Δhistidine), only yeast cells containing both AGB1-AGG1 and ARD1 were able to grow (Figure 3.3A). Yeast containing AGB1-AGG1 alone or ARD1 alone failed to grow, indicating that neither construct alone was responsible for the growth observed on Δhistidine media. These data suggest that ARD1 physically interacts with the AGB1-AGG1 heterodimer.

To test the validation of the Y3H results, we determined if recombinant ARD1 protein and AGB1-AGG1 heterodimer co-immunoprecipited from cell lysates. Lysed bacterial cells expressing a GST-tagged ARD1 protein were incubated with glutathione beads, and purified AGB1-AGG1 protein was added. The protein complexes were eluted from the beads and were subjected to SDS-PAGE. A polyclonal anti-AGB1 raised against an AGB1 peptide from the N-terminal helix of AGB1 was used (see Experimental Procedures) to detect the proteins by immunoblot analysis. As seen in Figure 3.3B,

AGB1 was co-immunoprecipitated with ARD1, corroborating the Y3H data and showing that these proteins interact physically *in vitro*.

In order to determine whether these proteins interact physically *in vivo* in *N. benthamiana* leaf epidermal cells, we utilized bimolecular fluorescence complementation (BiFC). Each bait and prey combination (ARD1 vs. AGB1, ARD1, or AHP2) was cloned into vectors containing either the C- or N-terminal half of YFP and all combinations were tested in tobacco leaves. A positive interaction was scored by the reconstitution of YFP as visualized under scanning confocal microscopy. ARD1 was able to reconstitute the fluorescent signal with AGB1, confirming the positive interaction between the two proteins (Figure 3.3C, top panel). Additionally, ARD1 was confirmed to form a homodimer or oligomer, which was shown for the rice protein OsARD1 (Sauter et al., 2005). As a negative control, we showed that ARD1 does not interact with a cytokinin-pathway cytosolic protein AHP2 (Figure 3.3C, bottom panel).

ARD1 has aci-reductone dioxygenase enzymatic activity that is stimulated by AGB1.

Recombinant 6x His-tagged ARD1 was purified in order to perform enzymatic assays. To determine which metal or metals were bound in the active site, ARD1 protein was analyzed using inductively coupled plasma mass spectrometry (Table 3.1). ARD in the on-pathway binds Fe^{+2} or Mg^{+2}, while ARD in the off pathway binds Ni^{+2}, Co^{+2}, or Mn^{+2} (Pochapsky et al., 2002). Therefore, we assayed for the presence of each of the five divalent metal ions. Fe(II) constituted 80% of the metal ions in the protein sample. The remaining ions in the sample were Mg^{+2} (13%), Co^{+2} (5%), and Ni^{+2} (2%). Together, these data suggest that ARD1 primarily performs the on-pathway reaction and that the

purified ARD1 displayed 100% metal occupancy. The equal molar concentration of Fe^{+2} and ARD1 indicate this metal is iron.

A poly-histidine tagged ARD1 protein was used to measure the enzymatic activity with aci-reductone. Because aci-reductone is short-lived and reactive with atmospheric oxygen, the reaction was performed anaerobically, and aci-reductone was generated *in situ* immediately prior to the addition of ARD1 using a previously published protocol (Zhang et al., 2004). ARD1 depleted the aci-reductone substrate, but no CO production was observed as detected by gas chromatography, suggesting that ARD1 does not significantly catalyze the off-pathway reaction. Although significant enzymatic activity was observed with a freshly prepared sample of ARD1 (Figure 3.4A open boxes, Figure 3.4B), its activity decreased over time, making it difficult to kinetically characterize the enzyme. After seven days at 4º C, all activity was lost; if the enzyme was stored at -20º C, the observed rates were 50% or less than the originally recorded numbers. Because of these factors, all rate comparison assays were performed on the same day and with the same batch of ARD1 that was used as a reference.

Because ARD1 and AGB1 interact physically (Figure 3.3), we tested whether AGB1 directly affected ARD1 enzymatic activity. We performed the same assay described above in the presence or absence of purified AGB1-AGG1 protein. In multiple experiments and independently of the ARD1 batch or the day on which the experiment was performed after its purification, AGB1-AGG1 consistently stimulated ARD1 activity about two-fold when mixed in a 1:1 ARD1:AGB1-AGG1 molar ratio (Figure 3.4A black boxes vs. open boxes, Figure 3.4B). The error was estimated based on the maximum error calculated in similar experiments. Although these experiments were repeated several

times, the instability of the enzyme and therefore the difference in initial rates from experiments performed on different days or with different preparations of enzymes made it impossible to directly compare rates from different experiments.

Identification of a potential AGB1-ARD1 interface.

In order to further characterize the interaction between AGB1 and ARD1, we sought to identify key residues that could form critical contacts at an interface between the two proteins. Because neither AGB1 nor ARD1 was crystallized alone or in complex, an evolutionary approach to predict binding regions between the two was necessary. For this analysis, our previously published prediction of novel Gβ interfaces based on the evolutionary history of Gβ proteins was used (Friedman et al., 2009). We identified surface regions of the AGB1 structure that are invariant between plants and mammals. This region was reduced by culling residues that formed a hydrogen bond with either the polypeptide backbone or with Gγ (which is required to stabilize Gβ) because some conserved surface residues are critical for structural maintenance, while others are required for protein-protein interactions that first evolved in plants and are maintained throughout eukaryotic evolution (see Figure 3.4A, "plant ancestor" in Friedman et al., 2009). Rather, residues with solvent-exposed functionality were assumed to be available for protein-protein interactions. Based on these criteria, a limited number of residues were chosen and four sets of corresponding mutant AGB1 proteins were created: the single mutations W109A and S129R, the double mutant E248K R25D, and the triple mutant Q120R T188K R235E (Figure 3.4C). W109 and S129 are both located in the conserved surface which is known from mammalian studies to bind Gα (Ford et al., 1998; Friedman et al., 2009). Additionally, W109 is an interaction "hot spot" in mammalian

proteins and is critical for interaction of Gβ with not only Gα but PLC-β2, AC2, and potassium and calcium channels (Ford et al., 1998). The remaining two regions had no previously identified functions in mammals, and thus represented key contacts of potential interaction interfaces with ARD1 (an interaction identified in plants but not yet identified in mammals). Mutant proteins were co-expressed with AGG1 (Gγ) containing a His tag for affinity purification and a mutated CAAX box for improved solubility. The AGB1-AGG1 heterodimer was purified from Hi5 insect cells by affinity column chromatography. The His-tagged AGG1 selected AGB1-AGG1 dimers (Figure 3.4D) assuring that the mutant AGB1 proteins were properly folded, as their interaction with AGG1 was necessary for purification. In addition, the mutant proteins physically interacted with ARD1 in a Y3H assay (Figure 3.4E) indicating that the mutant AGB1 proteins were still able to at least weakly bind to ARD1.

The mutated AGB1-AGG1 dimers were tested for their ability to activate ARD1 *in vitro*. The mutant S129R fully activated ARD1 to the same or greater level than wildtype AGB1-AGG1, suggesting that S129 is not critical for stimulation or that the S129R mutation confers a slight enhancement of AGB1-AGG1 stimulation of ARD1. The mutants W109A, E248K R25D, and Q120R T188K R235E abolished ARD1 stimulation, suggesting that these residues form critical contacts for ARD1 stimulation (Figure 3.4A, B). Together, these data suggest that these three groups of residues are required for AGB1-AGG1 to stimulate ARD1 enzymatic activity.

ARD1 activation mechanism.

We identified a region that was likely to both interact with Gβγ and to move in response to its binding, changing the accessibility of the ARD1 active site (Figure 3.5A).

The C-terminal α-helix (α-5) and preceding loop of ARD1 partially block the opening of the active site occupied by the metal ion. This obstruction likely lowers basal ARD1 activity; through displacement of this helix and/or loop, Gβγ may stimulate ARD1 activity. In addition, a conserved tryptophan residue exists in the middle of the loop, and is likely involved in the activation mechanism. In order to test the proposed mechanism of activation, we created and purified three versions of mutant ARD1 protein. One protein contained an alanine instead of the conserved tryptophan (W166A). The remaining proteins contained either a deletion of the C-terminal portion of α-5 (Δ188) or a deletion of all of α-5 and a portion of the preceding loop (Δ175). When tested in the absence of AGB1-AGG1 protein, only the W166A mutant had higher enzymatic activity than wildtype ARD1 (Figure 3.5B, C). However, when AGB1-AGG1 was added to the reaction, neither W166A nor Δ175 were stimulated by AGB1-AGG1 as seen in wildtype and in Δ188. These data suggest that W166 represses the enzymatic activity of ARD1 and that W166 and the region between residues 175 and 188 are required for stimulation by AGB1-AGG1. The C-terminal portion of α5 was not required for stimulation by AGB1-AGG1 and did not negatively regulate ARD1 enzymatic activity.

Discussion

AGB1 physically interacts with ARD1 to stimulate catalytic activity. Several lines of evidence support direct interaction and regulation, including site-directed mutagenesis to map roughly the protein-protein interface. The region of Gβγ that stimulates ARD1 activity is conserved between plants and mammals, raising the possibility that the activation mechanism is ancient.

ARD proteins are present in bacterial species while heterotrimeric G protein components are not. How is it possible that Gβ proteins stimulate ARD enzymatic activity in eukaryotes but that this stimulation is not required for prokaryotic activity? In order to answer this question, we created a model predicting the molecular mechanism of activation of ARD1 by AGB1-AGG1 (Figure 3.5). The three ARD1 binding regions on AGB1 are spread throughout the top surface of AGB1, and the surface area is large enough that binding to the whole region would require a conformational change of ARD1. Upon examination of the modeled structure of ARD1 (Figure 3.S1) and of its parent structure (mADI1, PDB ID 1VR3 (Xu et al., 2006)), we noticed several features that would allow such an interaction to occur. First, the C-terminal α-5 helix of ARD1 and the loop directly preceding it pack tightly against the opening to the pocket that the metal ion and substrate occupy, obscuring the opening. We propose that this packing creates a low basal activity of ARD1. Binding of AGB1-AGG1 would displace the helix and loop, opening the ARD1 pocket and allowing for faster substrate/product movement into and out of the active site. Furthermore, the large, hydrophobic tryptophan residue in the middle of the loop appears to regulate this movement.

We tested these predictions by creating two truncation mutations (Δ175 and Δ188) and one substitution mutation (W166A). The higher basal activity of W166A indicates that W166 restricts the basal enzymatic activity of ARD1. However, it is also a critical point of regulation by AGB1-AGG1, as AGB1-AGG1 was unable to further stimulate this mutated protein. The portion of the loop and α-5 helix between residues 175-188 did not restrict basal ARD1 activity but was critical for stimulation by AGB1-AGG1 (the C-terminus of α-5 was not involved in this process). Therefore, we suggest that W166 serves as a latch that maintains ARD1 in a state of low basal enzymatic activity ("inactive"). Upon binding by AGB1-AGG1, this latch is "unlocked" and the enzyme's activity is stimulated. AGB1-AGG1 utilizes both W166 and at least a subset of the residues 175-188 to achieve this stimulation.

These predictions are supported by several structural observations. First, while the crystallized mammalian structure was in the "inactive" conformation, with the loop and C-terminal helix obscuring the pocket, the Ni-ARD bacterial structure (PDB ID 1ZRR (Pochapsky et al., 2006)) showed a displaced loop and helix and a much more accessible pocket, and the Fe-ARD bacterial structure (PDB ID 2HJI (Ju et al., 2006)) showed a disordered C-terminius that indicates mobility of the C-terminal loop and helix. Thus, the structural differences observed agree with our proposed method of stimulation; since the bacterial structures exist in a more "active" state, their only regulation may be from the divalent metal bound the active site. However, the "inactive" state of the eukaryotic protein suggests the need for regulation by an outside protein, in this case Gβγ.

The best understood function of ARD proteins is to operate in the methionine salvage pathway, utilizing one of two classes of divalent metal ions to affect the structure and thus to convert aci-reductone into either a methionine precursor or an off-pathway product. The high sequence similarity between ARD1 and the previously characterized KoARD and MmADI1 within the regions required for enzymatic activity (Figure 3.S3A) suggested that ARD1 might function in the methionine salvage pathway. Our results confirmed this prediction. However, it is also known that mammalian ARD enzymes are "moonlighting" proteins (Jeffery, 2003), performing diverse functions in addition to (and independently of) the enzymatic activity. For example, human ARD (hADI1) serves as a cancer suppressor in prostate cells. When upregulated in tumor cells, hADI was found to induce programmed cell death independent of its enzymatic function (Oram et al., 2007). hADI1 has also been shown to move between the cytoplasm and nucleus via an undefined nuclear localization signal and a non-canonical nuclear export signal (NES). One of its nuclear functions is to facilitate mRNA splicing; this function is also independent of methionine salvage enzymatic activity (Gotoh et al., 2007).

ard1 null mutants have shorter hypocotyls than wildtype due to fewer cells but the reason is unresolved. This may involve intermediates of the methionine pathway or some other function of ARD1 (discussed further below). The methionine salvage pathway recovers the thiomethyl group of methylthioadenosine (MTA, Fig 1E), a byproduct of polyamine biosynthesis, and it has been shown that polyamines are critical for cell division and cell cycle maintenance (Oredsson, 2003). In addition, MTA, a byproduct of the methionine salvage cycle, inhibits polyamine synthesis via spermine synthase and spermidine synthase (Avila et al., 2004). Previous studies showed that mutants of several

methionine salvage pathway components displayed altered polyamine levels (Kim et al., 2008; Burstenbinder et al., 2010). Therefore, it is possible that a decrease in the polyamine pool contributes to the reduced cell division observed in *ard1* and *agb1-2* hypocotyls.

Another hallmark of two-day-old etiolated seedlings is a closed apical hook; this phenotype protects the apical meristem from damage as the seedling grows through the soil, and ethylene production contributes to the maintenance of the hook (Woeste and Kieber, 1998). Because two-day-old etiolated *ard1* and *agb1-2* seedlings display an opened apical hook, we speculate that the decrease in ethylene production in these mutants may contribute to this tissue-specific phenotype. It is important to note that the decrease in ethylene production measured at this developmental stage is, while statistically significant, quite small (15-25%) although we are aware that such a small change in ethylene has been correlated to a significant alteration in hypocotyl length (Tsuchisaka et al., 2009).

Arabidopsis and other plants (but no non-plant organisms) contain multiple ARD proteins; the *Arabidopsis* genome encodes AtARD1-4. AtARD1 is 96% identical to AtARD2 and 80% identical to AtARD3 but only 67% identical to AtARD4. These *Arabidopsis* ARD orthologs are more similar to ARD proteins from rice than to ARDs from non-plant species, but the ARD1, ARD2, and ARD3 of different plants are more closely related to each other (between species) than to the ARD4 of their own species (Figure 3.S3B). Each of the four *Arabidopsis* ARD proteins can interact with Gβγ *in vitro* and *in vivo* (Figure 3.S4); G protein signaling may also utilize these ARD paralogs.

Although there is a high degree of similarity between ARD1-3, ARD4 is divergent and we speculate could possess a different subset of functions than the other three. The duplication of ARD proteins in plants is of particular interest in that it occurred after the divergence of plants from the ancestral lineage but did not occur in any other family. Additionally, ARD1 likely diverged from ARD4 before the divergence of rice and *Arabidopsis* and then duplicated twice to yield ARD1, ARD2, and ARD3 after the split (Figure 3.S3B). This pattern and lack of subsequent divergence indicates a functional necessity for multiple ARD family members. Though plant ARD proteins display conserved metal-binding residues, suggesting enzymatic activity, *Arabidopsis* ARD proteins also possess the conserved non-canonical NES identified in the mammalian ARD proteins (Figure 3.S1B, underlined region), suggesting a nuclear role as well.

It is still not yet clear what other processes ARD proteins might regulate in plants and which (if any) of these additional functions are also regulated by G protein signaling. We identified a role of ARD1 in cell division and that role depends on the ability of AGB1 to stimulate ARD1 enzymatic activity. Here, one clear picture of regulation of ARD1 activity by AGB1 emerges, but other functions and mechanisms should likely follow.

Acknowledgements

We thank Arwen Frick-Cheng for technical assistance in creating the baculoviruses used to generate the various forms of AGB1 protein and Gyeong Mee Yoon and Joe Kieber for the use of their gas chromatograph and advice in measuring ethylene production. Work in A.M.J.'s lab on the *Arabidopsis* G proteins is supported by

grants from the NIGMS (R01GM065989), DOE (DE-FG02-05er15671) and NSF (MCB-0723515 and MCB -0718202). Work in T.C.P.'s lab is supported by a grant from the NIH (R01-GM44191).

Materials and Methods

Screen for suppressors of agb1-2 and plasmid rescue.

sgb3-1^{D} was identified in the activation tagging suppressor screen that was previously described by Wang et al. (Wang et al., 2006).

Plant materials.

All *Arabidopsis* seeds were in the Columbia-0 ecotype. The following T-DNA insertion mutant alleles were used: *ard1-1* (SALK_119327), *ard1-2* (GABI_595C04) (Rosso et al., 2003), and *ard1-3* (SALK_034308). T-DNA insertions were confirmed by full-length genotyping PCR using primers listed in Table 3.S1. mRNA transcript levels were identified by RT-PCR. Total RNA was isolated from rosette leaves using the RNEasy plant mini kit™ (Qiagen, Valencia, CA), and cDNA synthesis was performed using a poly(dT) primer. RT-PCR was performed with the SuperScript III RT-PCR kit™ (Invitrogen, Carlsbad, CA). PCR was performed for 25 cycles (94ºC for 30 sec, 55ºC for 30 sec, 72ºC for 1 min) using the primers listed in Table 3.S1. *agb1-2* has been previously described (Ullah et al., 2003). Gene accession numbers are as follows: AGB1, At4g34460; ARD1, At4g14716; ARD2, At4g14710; ARD3, At2g24600; ARD4, At5g43850.

Hypocotyl and hook assays and epidermal cell counting.

Seeds were sterilized in 70% ethanol for 1 min, 30% bleach plus 0.01% Tween-20 for 20 min, and rinsed 3-5 times with sterile distilled water. Seeds were sown on square Petri plates containing ½ Murashige & Skoog (MS) salts, 1% sucrose, and 0.6% phytagel. The plates were incubated at 4 ºC in the dark for 2-4 d, exposed to light for 2h, and were grown vertically in the dark for 2d (52h) or 4d at 22ºC. Upon opening, the plates were

immediately scanned and imaged. Apical hook angles and hypocotyl lengths were quantified using ImageJ software. For epidermal cell counting, 52h seedlings were incubated overnight in a solution of 8:2:1 chloral hydrate (1mg/ml), glycerol, and water. Epidermal cells were imaged using Nomarski optics as described previously (Ullah et al., 2001).

Ethylene biosynthesis.

Seeds were sterilized as reported above and sown in a 0.06% agarose suspension in 22 ml gas chromatography vials containing 3ml of 1X MS salts, 1% sucrose, and 0.6% agar. 100-200 seeds were sown for ethylene measurements at 2d. The vials were placed at 4ºC in the dark for 2-4 d, exposed to light for 2h at 22ºC, and then capped and grown in the dark at 22ºC for 2d. Ethylene was measured by gas chromatography as previously described (Vogel et al., 1998). At least three vials were measured for each genotype. The experiment was repeated once.

Statistics.

Statistical analyses for the plant growth measurements were performed using a type three two-tailed Student's t-test.

Phylogenies and Bioinformatics.

Representative ARD sequences were collected via a BLAST analysis across the species indicated using ARD1 as a query. The alignment was generated using ClustalX (Thompson et al., 1997). MrBayes (Huelsenbeck et al., 2001) was run using a fixed equalin model, using the inverse gamma rate, and sampling 1,000,000 generations at a frequency of 100 for 3 independent runs with a burn in of 250,000 generations to generate a consensus phylogenetic tree.

Modeling.

A homology model of ARD1 was created using MmADI1 (PDB ID 1VR3) as a model; this structure was identified by a BLAST search and chosen for its high sequence similarity. The model was created using the Insight II software program from Accelrys, Inc (San Diego, CA). The homology model was evaluated with the Verify-3D function of Insight II, and the resulting normalized score was 0.74, indicating a viable structure.

ARD1-GFP transient expression in Arabidopsis protoplasts.

ARD1-GFP The coding region of ARD1 was cloned into the GFP expression vector pB7FWG2. *Arabidopsis* mesophyll protoplasts were generated from wildtype and *agb1-2* plants as previously described (Yoo et al., 2007). Cells were imaged using a Zeiss LSM 710 confocal laser scanning microscope equipped with a C-Apochromat X40 (NA 1.2) water immersion objective and standard Zeiss software (ZEN).

Yeast-Three Hybrid (Y3H) Protein Interaction.

AGB1 and AGG1 were cloned into the pBridge^{TM} vector (Clontech, Palo Alto, CA). ARD1 was cloned into the p-ENTR/D-TOPO^{TM} vector (Invitrogen, Carlsbad, CA) and then recombined into the pACTGW-attR Gateway vector which contains an activation domain and is compatible with the pBridge^{TM} vector. The bait (AGB1-AGG1) and prey (ARD1) were transformed separately or together into yeast strain AH109. All three strains (bait alone, prey alone, or bait and prey) were grown on nutritionally selective media. The presence of the bait and prey was confirmed by the expression of nutritional markers (positive growth on media lacking tryptophan and leucine, respectively). Interaction was confirmed by the expression of an additional nutritional marker (positive growth on media lacking histidine). For Y3H experiments involving

AGB1 mutants, the coding sequence of wildtype AGB1 and of each mutant was cloned into the pAS vector, which is compatible with the ARD1-pACTGW-attr vector.

Protein Expression (ARDs, AGB1, used for pull-downs and enzymatic assays).

The coding regions of ARD1-4 were cloned into pDEST17™ and pDEST15™ (Invitrogen, Carlsbad, CA) containing an N-terminal poly-His tag and an N-terminal GST tag, respectively, and expressed in BL21-rosetta *E. coli* cells. Protein expression was induced at OD=0.6 for 2h at 37ºC. ARD-His was purified on a Talon column (Clontech, Mountain View, CA). ARD-GST was purified on an immobilized glutathione column (Pierce, Rockford, IL). Protein concentrations were determined using the Coomassie Plus Reagent™ (Pierce, Rockford, IL). To purify Gβγ protein from insect cells, AGB1 was cloned into pDEST8 (untagged vector), and AGG1 containing a C-terminal mutation of the CAAX box prenylation site (C95*) was cloned into pDEST10™ (N-terminal 6xHis tag) (Invitrogen, Carlsbad, CA). Baculoviruses against each protein were generated and propagated using the Bac-to-Bac Baculovirus Expression System (Invitrogen, Carlsbad, CA). Viruses were propagated separately in Sf9 cells, were co-expressed in Hi-5 cells, were purified on a Talon™ column (Clontech, Mountain View, CA), and concentrations were determined as described above.

AGB1 polyclonal antibody.

A peptide consisting of 18 N-terminal residues (T14-L31) of AGB1 was synthesized and conjugated to the keyhole limpet hemocyanin. This peptide-carrier was used to raise an antibody in rabbits (Open Biosystems, Huntsville, AL). In total, five immunizations were administered. The terminal serum was incubated at a dilution of

1:20,000 in phosphate buffered saline containing 5% milk and 0.01% Tween-20 for 2h to detect purified proteins. The antiserum lot used here was designated AGB1-1607-Ext.

Co-imunoprecipitation.

BL21 bacterial cells expressing a GST-tagged ARD protein (ARD1, 2, 3, or 4) were lysed (sonicated in 25mM Na-P, 300 mM NaCl, 0.2mM PMSF, and 1x protease inhibitor cocktail (Sigma, St. Louis, MO)) and then incubated with glutathione beads. After washing, purified AGB1-AGG1 protein was added to the ARD-bead mixture. After incubation, the slurry was washed several times, and the protein complexes were eluted from the beads by boiling in SDS loading buffer. A subset of the elution product was run on an SDS-PAGE gel. The proteins were transferred to a nitrocellulose membrane and probed with the polyclonal AGB1 antibody described above.

Bimolecular Fluorescence Complementation (BiFC).

BiFC was performed according to Grigston et al. (Grigston et al., 2008) with several modifications. The coding sequences of ARD1, ARD2, ARD3, ARD4, and AHP2 were cloned into the BiFC vectors pCL112_JO (YFP-n) and pCL113_JO (YFP-c). The coding sequence of AGB1 was cloned into pBatL-sYFP-C (YFP-c). Constructs were additionally co-infiltrated with mitochondrial RFP (mt-Rk obtained from the Arabidopsis Biological Resource Center, CD3-991) as a positive transformation control. Samples were imaged using a Zeiss LSM 710 confocal laser scanning microscope equipped with a C-Apochromat X40 (NA 1.2) water immersion objective and standard Zeiss software (ZEN). Visualization of YFP was achieved using a 514 nm Argon laser line for excitation and the photomultiplier detector was set to collect emission bandwidth at 526-569 nm.

Visualization of RFP was achieved using a 560-nm diode laser for excitation and the photomultiplier detector was set to collect emission bandwidth at 565-621 nm.

AGB1 and ARD1 mutagenesis.

The coding sequences of AGB1 and ARD1 were mutagenized using the QuickChange Multi and QuickChange Lightning Site-Directed Mutagenesis Kits™, respectively (Agilent Technologies, Santa Clara, CA). Mutagenesis primers were designed using the QuickChange Primer Design Program. Mutations were confirmed by sequencing and were expressed and purified as described above.

Inductively Coupled Plasma Mass Spectroscopy (ICP-MS).

Measurements were performed in the Mass Spectrometry Facility of the Department of Chemistry, UNC-Chapel Hill (Chapel Hill, NC) by using a Varian 820 ICP-MS (Palo Alto, CA), now part of Bruker Daltonics, (Billerica, MA). Protein samples were prepared by adding 1 mL 70% HNO_3 to 40 μg of protein (sample 1) and 24 μg of protein (sample 2) for ~10 hours. Each sample was then diluted with 2% HNO3 solution to a final volume of 10mL. All solutions were made with 18 mega-ohm deionized water. A 5 ppb indium solution (in 2% HNO_3) was used as the instrument internal standard during the runs. Isotopes ^{24}Mg, ^{55}Mn, ^{57}Fe, ^{59}Co, and ^{60}Ni were quantitatively monitored in a peak hopping mode with a total of 100 scans per point per isotope (50,000 μsec of dwell time, and 5 replicates of 20 scans).

Carbon Monoxide Assay.

The carbon monoxide (CO) assay was performed according to Sundin and Larsson (Sundin and Larsson, 2002). The formation of CO was monitored and quantified using gas chromatography (GC). A Hewlett-Packard 6890 GC system equipped with an

Agilent G2747A nickel catalyst system with a flame ionization detector was used for the experiment. The injector temperature was set at 60ºC, the oven temperature at 35ºC, the nickel catalyst temperature at 375ºC and the detector temperature at 250ºC. The enzyme catalyzed solution was placed in a gas tight sealed vial with 1 ml CO liberating solution (7.5 g saponin in 1 M sulfuric acid). This vial was then vortexed for 1 min and shaken for 40 min at 250 rpm at 37º C. An aliquot of 100 µL from the headspace gas phase was injected with a gas-tight syringe into the inlet of the GC column.

Enzymatic Assay.

The enzymatic assays to probe for activity of ARD1 and mutant proteins were performed according to Zhang et al. (Zhang et al., 2004) with some modifications. All experiments were performed in three consecutive steps in an anaerobic cuvette. First, the substrate aci-reductone was built to the final concentration of about 125 µM by 75 nM E1 enzyme and in the presence of 200 µg/ml catalase. Then, buffer saturated with molecular oxygen (280 mM) was added, and the rate of aci-reductone decay was measured at 308 nm. The average oxygen-induced decay rate was $8.5\times10^{-11}\pm1.5\times10^{-11}$mol substrate/second. This rate was accounted for in the calculations of the initial rates of the enzymes. Lastly, a controlled amount of ARD1 variant was added, and the depletion of aci-reductone was monitored at 308 nm for at least 300 s. The initial rates were calculated by selecting the linear portion of the graph and calculating the linear fit in this region.

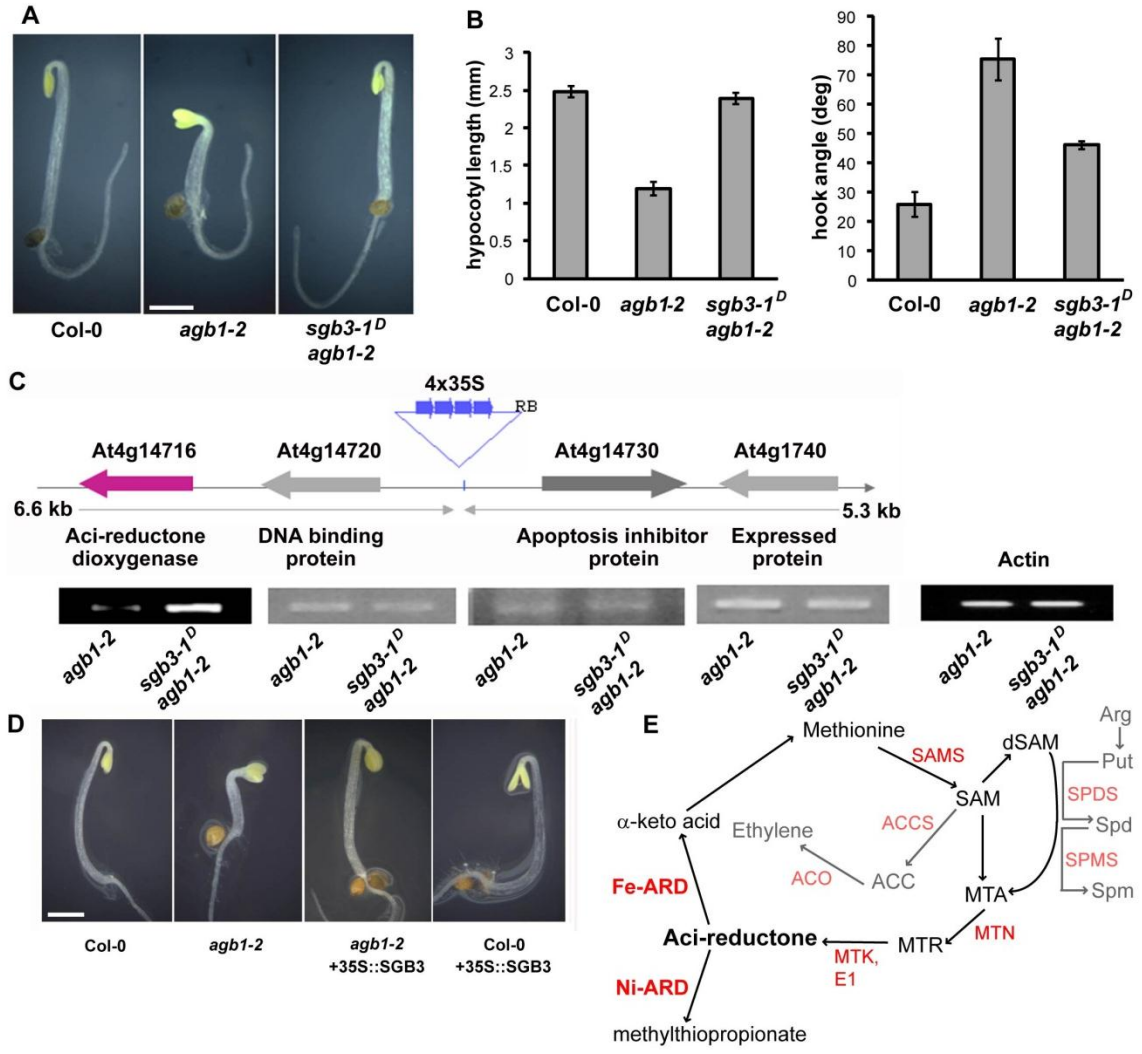


Figure 3.1. ARD1 is a genetic suppressor of *agb1-2*. *A.* Two-day-old etiolated phenotype of wild-type, *agb1-2*, and *sgb3-1*^{D}/*agb1-2* seedlings. Bar = 1 mm. *B.* Quantification of the hypocotyl lengths and apical hook angles observed in (A). Error bars = standard error. *C.* Enhancer locus identification and confirmation of enhanced SGB3 transcript levels by RT-PCR. PCR products were run on two separate gels in non-adjacent lanes. *D.* Recapitulation of the *sgb3-1*^{D}/*agb1-2* phenotype by overexpressing *SGB3* in the presence and absence of *AGB1*. *E.* The methionine salvage pathway as defined by experiments in bacterial and plant systems (Ju et al., 2006; Burstenbinder et al., 2010). SAM is recycled into methionine via the intermediates MTA, MTR, and aci-reductone. Ni-ARD catalyzes an off-pathway reaction while Fe-ARD catalyzes an on-pathway reaction and promotes the recycling of methionine. Characterized enzymes are indicated in red (SAMS, SAM synthase; MTN, MTA nucleosidase; MTK, MTA kinase; E1, E1 enolase/phosphatase). Ethylene is produced from SAM via 1-aminocycloprane-1-carboxylate (ACC) by ACC synthase (ACS) and ACC oxidase (ACO). SAM is also decarboxylated (dSAM) and facilitates the synthesis of the polyamines spermidine (Spd, via Spd synthases, SPDS) and spermine (Spm, via Spm synthases, SPMS) from putrescin (Put).

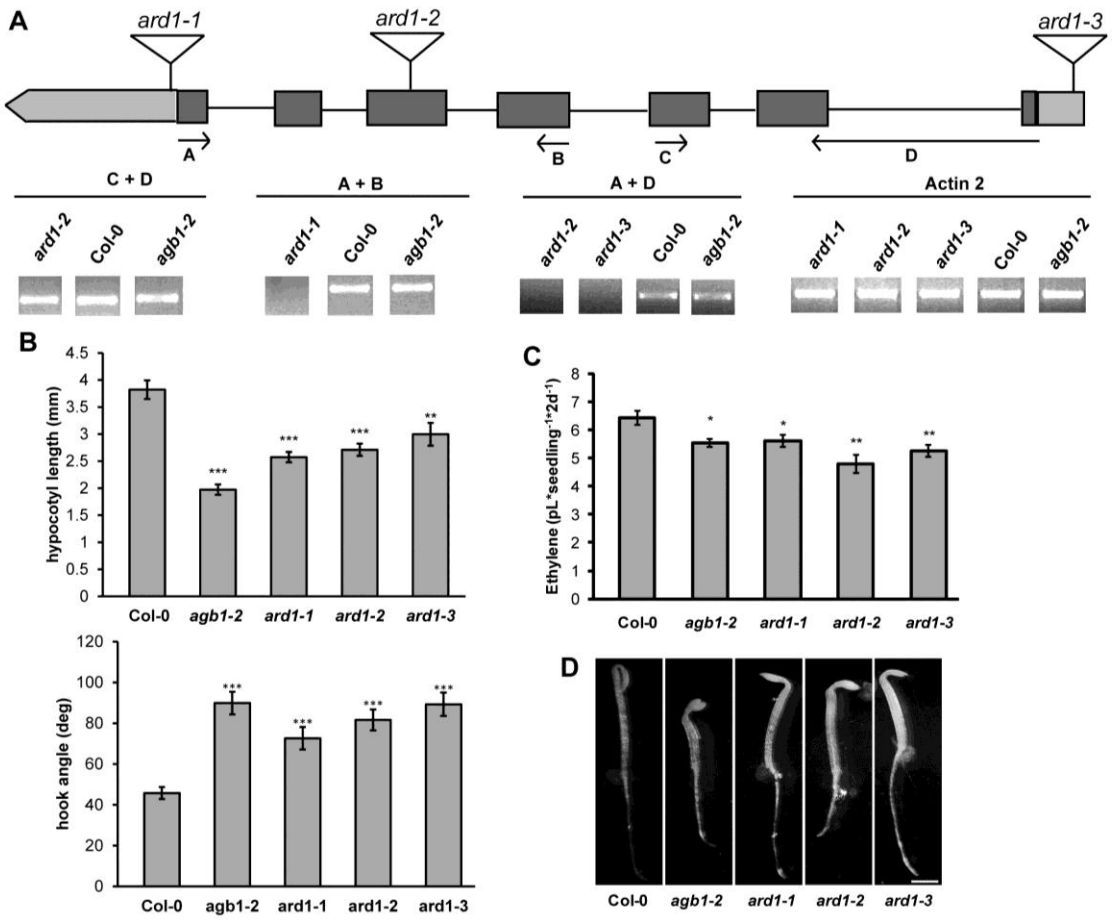


Figure 3.2. ARD1 null mutations affect cell division and ethylene production. *A*. Location of three T-DNA insertions in *ARD1*. Arrows indicate genotyping primers used to amplify the mRNA transcript. RT-PCR was performed using the primers indicated. *Actin2* was used as a reference transcript. PCR products were run on two separate gels in non-adjacent lanes. *B*. Hypocotyl lengths and apical hook angles of two-day-old etiolated seedlings. *p<0.05, **p<0.005, ***p<0.0005. Results were an average of two independent experiments, and are representative of three additional experiments performed. Error bars = standard error of the mean. *C*. Ethylene biosynthesis of wildtype and *ard1* or *agb1* mutant plants grown in the dark for 2d. Each measurement is an average of at least three vials, and the experiments were repeated with similar results. *p<0.05, **p<0.005. Error bars = standard error. *D*. Etiolated phenotype of two-day-old seedlings. Bar = 1 mm.

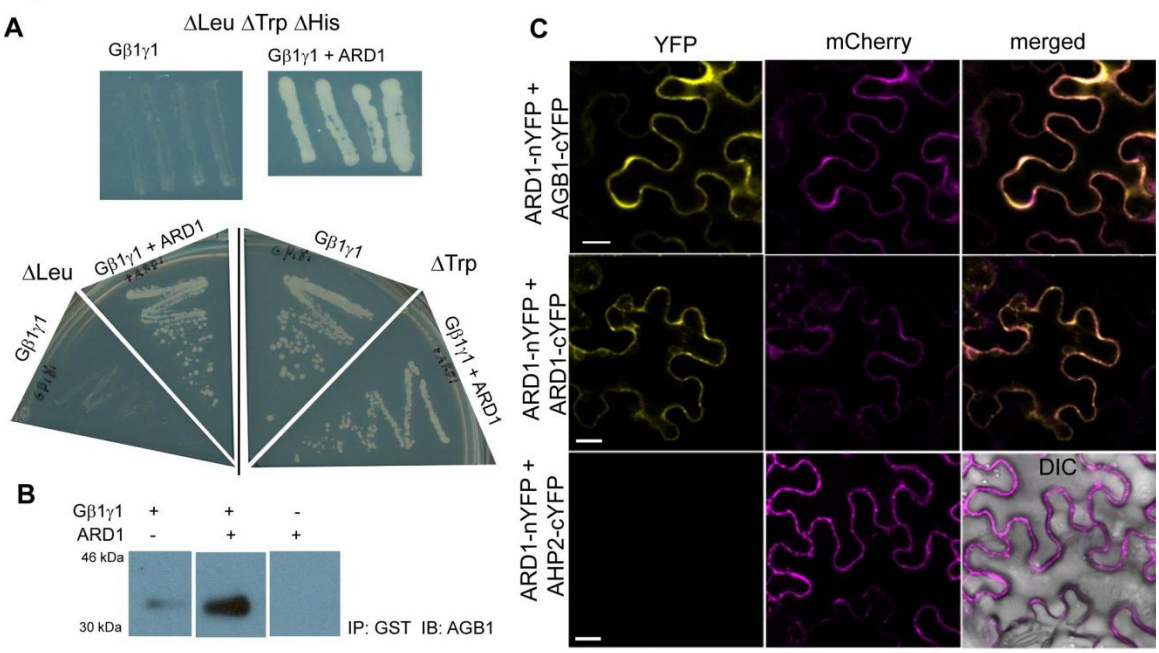
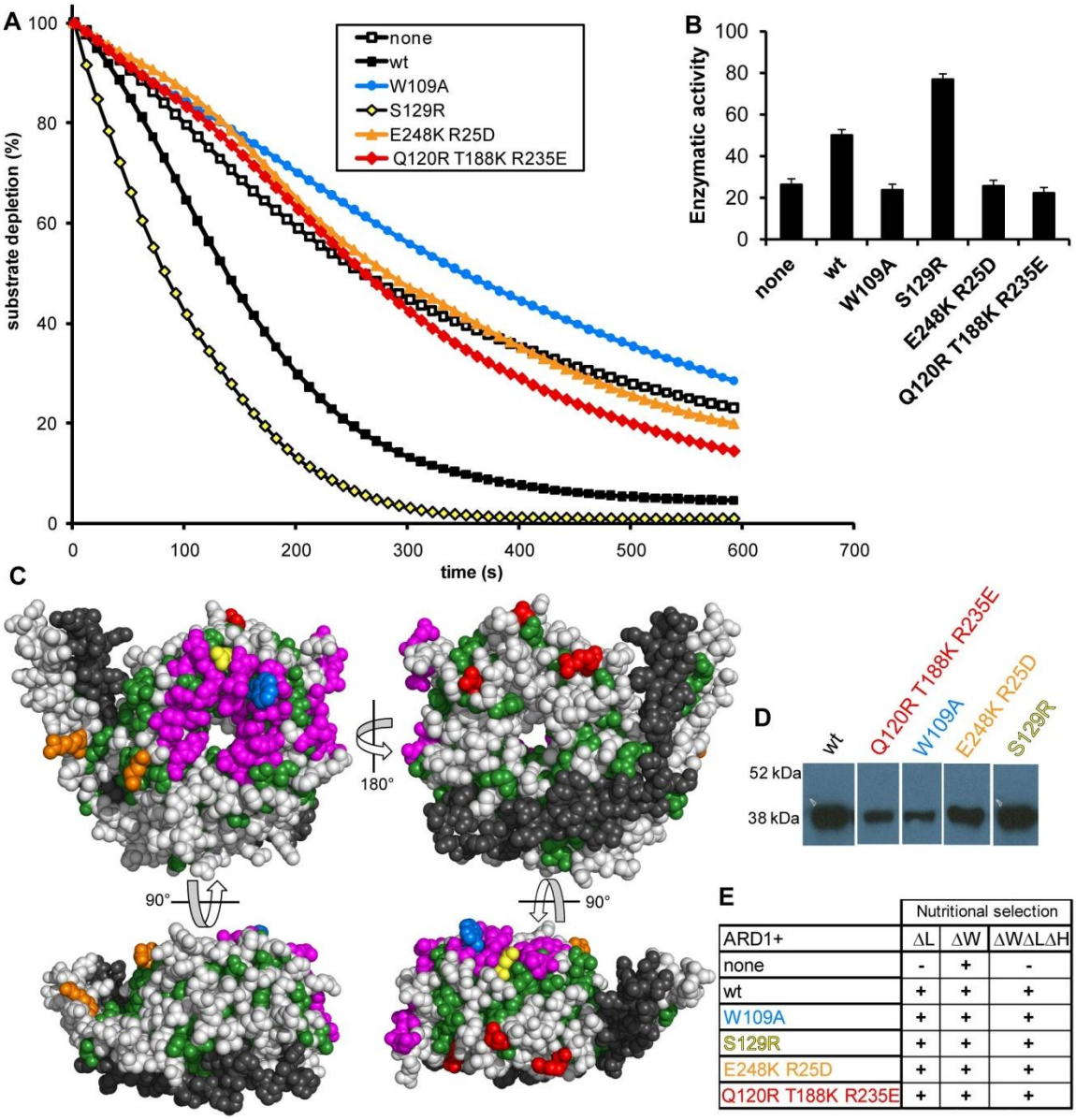


Figure 3.3. AGB1 and ARD1 interact physically. *A.* Growth of yeast strain AH109 containing the proteins indicated (AGB1-AGG1 (AtGβγ) alone or AtGβγ + ARD1) on yeast dropout media missing leucine, tryptophan, and histidine. This selects for a positive interaction between each of the two genes, resulting in no growth for the strain containing AtGβγ alone and positive growth for the strains containing both AtGβγ and ARD1. Four independent colonies were restreaked on the triple dropout media to confirm the interaction. Positive growth on media lacking leucine or tryptophan confirms the presence of ARD1 or AtGβγ, respectively. *B.* 6x-His-tagged Gβγ was immunoprecipated with ARD1-GST on a glutathione resin and detected by immunoblotting with anti-Gβγ. Proteins were separated on one gel in non-adjacent lanes *C.* Bifluorescence molecular complementation in *N. benthamiana* leaf epidermal cells (error bar = 20 μM). ARD1 interacts with AGB1and homodimerizes. ARD1 does not interact with cytosolic protein AHP2 (negative control, merged panel includes DIC image to show the cell borders). Mitochondrial RFP expression (mCherry) confirms positive transformation of all cells.



E

	Nutritional selection		
ARD1+	ΔL	ΔW	ΔWΔLΔH
none	-	+	-
wt	+	+	+
W109A	+	+	+
S129R	+	+	+
E248K R25D	+	+	+
Q120R T188K R235E	+	+	+

Figure 3.4. AGB1-AGG1 stimulates ARD1, and this stimulation is reduced in several AGB1 mutants. *A.* ARD1 enzymatic activity in the presence and absence of wildtype AGB1-AGG1 (Gβγ) and various AGB1 mutants in a 1:1 molar ratio. *B.* Enzymatic rates of ARD1 activity in the presence of each AGB1-AGG1 wildtype or mutant construct indicated ("none" indicates ARD1 alone). The rates are expressed in moles of substrate per mole of enzyme per second +/- standard deviation and were recorded as initial rates. These rates account for the average oxygen-induced decay rate (baseline, see Experimental Procedures). *C.* Four views of the AGB1 protein surface. All colored residues compose a region strictly conserved between plant and mammalian species. Magenta residues form the Gα binding interface. Colored residues indicate point mutations created, as indicated in (D). *(continued on next page)*

Figure 3.4 *(continued)*
D. AGB1 mutant proteins were co-expressed in *E. coli* with AGG1 and purified via the His tag on AGG1 by affinity column chromatography (IB: anti-AGB1). Proteins were separated on one gel in non-adjacent lanes. *E.* Y3H growth of strains containing ARD1 in the presence or absence of wildtype and mutant AGB1 proteins. Cells expressing the genes encoding mutant AGB1 proteins were able to grow on media lacking histidine, indicating that the mutations do not disrupt the physical interaction between AGB1 and ARD1.

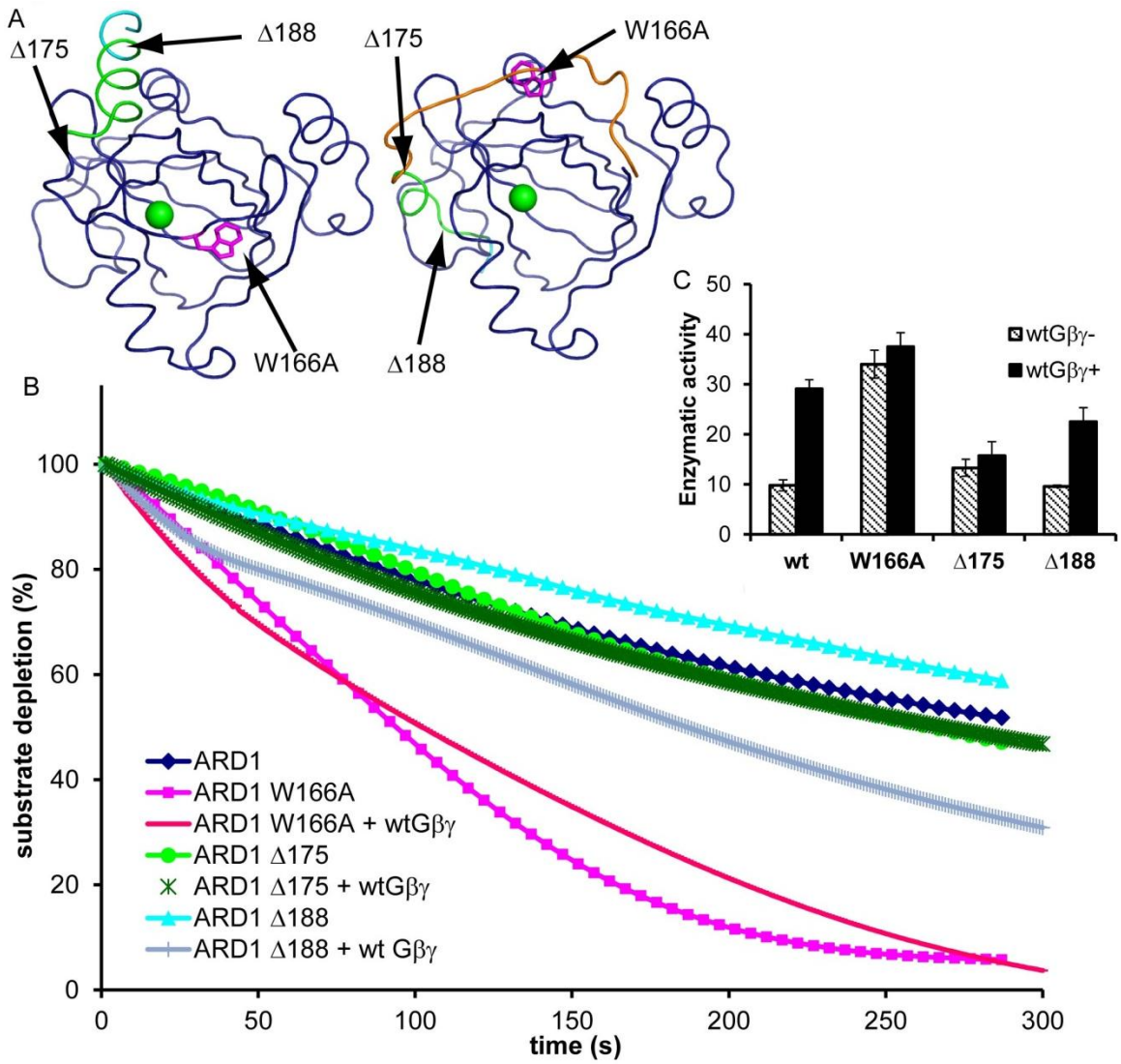


Figure 3.5. Mechanism of ARD1 stimulation by AGB1. *A.* Structural model of ARD1 depicting the proposed mechanism of stimulation by AGB1-AGG1. Green sphere represents the bound metal ion in the active site. Blue and orange loop and helix represent the C-terminal portion of the protein in the inactive and active conformation, respectively. The three mutations made on ARD1 are indicated in the structure. *B.* Enzyme activity of wildtype vs. mutant ARD1 in the presence or absence of wildtype AGB1-AGG1 (Gβγ). The final enzyme concentration was adjusted to 24 nM and the substrate was built up to about 125 μM in all cases. The molar ratio of ARD1:Gβγ was 0.65:1 *C.* Enzymatic rates of ARD1 wildtype and mutants in the presence or absence of wildtype Gβγ as indicated in (B). The rates are expressed in moles of substrate per mole of enzyme per second +/- standard deviation and were recorded as initial rates on the third day after purification. wtGβγ- indicates the rate for the enzyme alone, and wt Gβγ+ indicates the rate of the enzyme with the addition of wildtype Gβγ. These rates account for the average oxygen-induced decay rate (baseline, see Experimental Procedures).

Table 3.1. Quantification of metals bound to ARD1. Inductively coupled plasma mass spectrometry analysis for the presence of ^{24}Mg, ^{55}Mn, ^{57}Fe, ^{59}Co, and ^{60}Ni. Elemental concentration is expressed as parts per billion (ppb).

Sample	**Elemental Concentration (ppb)**
EF-01	
Mg24	1.56
Mn55	0.00
Fe57	9.34
Co59	0.63
Ni60	0.21
EF-02	
Mg24	0.62
Mn55	0.00
Fe57	7.26
Co59	0.44
Ni60	0.08

Supplementary material

The following figures were submitted to JBC as supplementary materials. Figure 3.S1 highlights the features of the ARD1 sequence and structure. Figure 3.S2 shows the subcellular localization of ARD1 in the presence and absence of AGB1. Figure 3.S3 shows the relationship between ARD1 and other ARD proteins, and Figure 3.S4 shows the interactions between the four ARD proteins and AGB1.

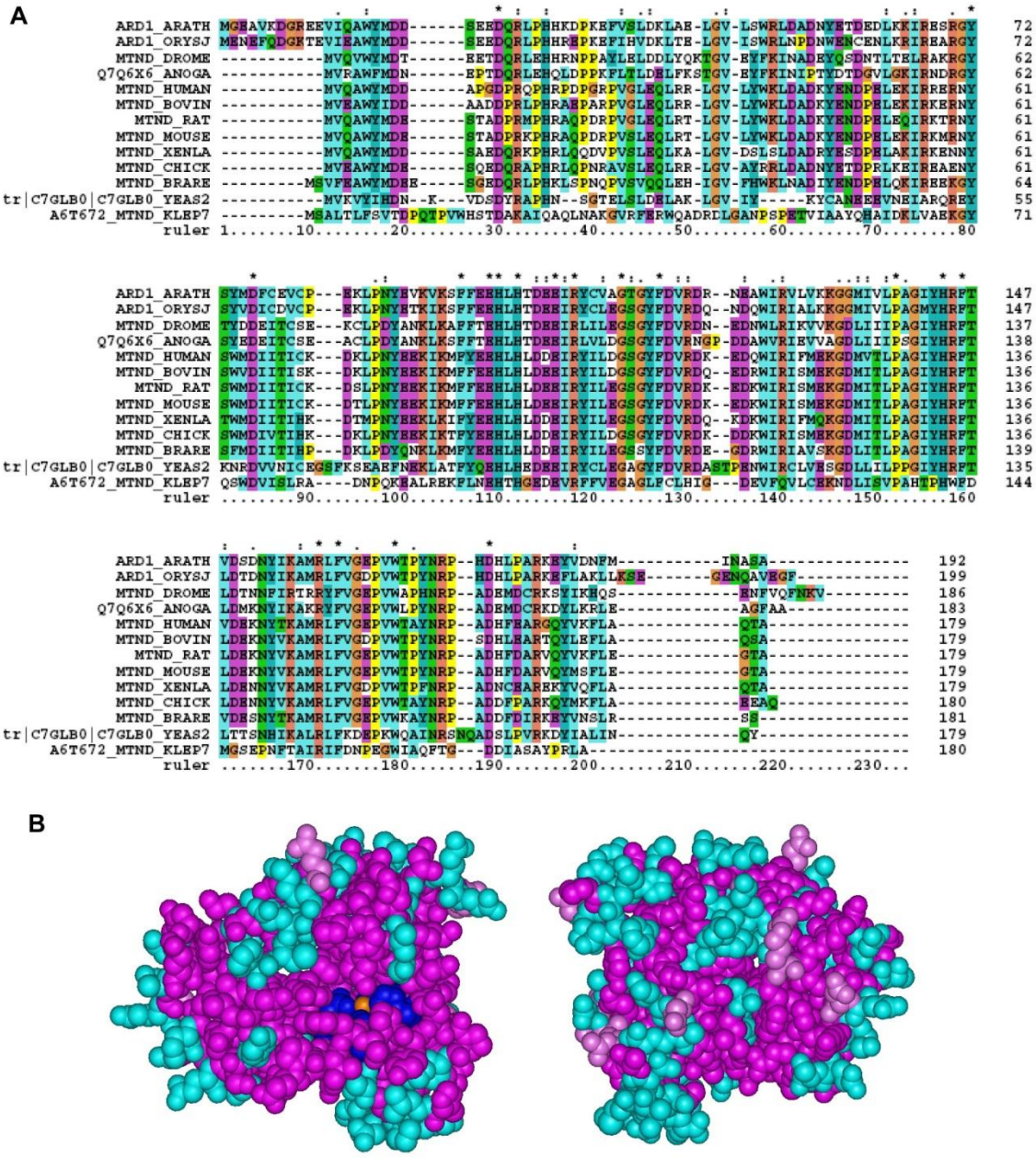


Figure 3.S1. ARD1 is an aci-reductone dioxygenase protein. *A.* Multiple sequence alignment of ARD proteins (ARATH, Arabidopsis; ORYSJ, rice; DROME, drosophila; ANOGA, mosquito; HUMAN, human; BOVIN, bovine; RAT, rat; MOUSE, mouse; XENLA, xenopus; CHICK, chicken; BRARE, fish; YEAS2, yeast; KLEP7, bacteria). Circled star indicates four conserved metal-binding residues. *B.* Homology model of ARD1 threaded into the crystal structure of MmADI1 (PDB ID 1VR3, homology model score = 0.74). Magenta residues are conserved between the two, while cyan residues are plant-specific.

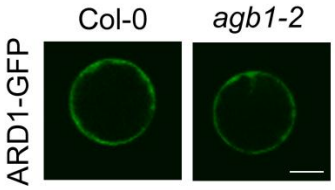


Figure 3.S2. ARD1-GFP expression in wildtype and *agb1-2 Arabidopsis* protoplasts. Bar = 10 μM.

A

```
ARD1      MGEAVKDGREEVIQAWYMDDSEEDQRLPHHKDPKEFVSLDKLAELGVLSWRLDADNYETD  60
ARD3      MGEAAKDQTEEVIQAWYLDNKEEDQKLPHHKDPKEFVSLDKLAELGVLCWRLDADNYETD  60
ARD2      MGEVVKDGREEVIQAWYMDDSEEDQRLPHHKDPKEFLSLDKLAELGVLSWRLDADNYETD  60
ARD4      ----------MALEAWFMDDSNEDQRLPHHRNPKELVSLDYLAELGVLYWKLNPENYEND  50
                     .::**::*:.:***:****::***::*** ******* *:*:.:***.*

ARD1      EDLKKIRESRGYSYMDFCEVCPEKLPNYEVKVKSFFEEHLHTDEEIRYCVAGTGYFDVRD 120
ARD3      EELKRIRESRGYSYMDLCEVCPEKLPNYEEKVKMFFEEHLHIDEEIRYCLAGSGYFDVRD 120
ARD2      EDLKKIRESRGYSYMDFCEVCPEKLPNYEVKVKSFFEEHLHTDEEIRYCVAGSGYFDVRD 120
ARD4      SELSKIREDRGYDYMDLLDLCPEKVSNYEEKLKNFFTEHIHKDEEIRYCLAGSGYFDVRD 110
          .:*.:***.***.***: ::****:.*** *:* ** **:* *******:**:*******

ARD1      RNEAWIRVLVKKGGMIVLPAGIYHRFTVDSDNYIKAMRLFVGEPVWTPYNRPHDHLPARK 180
ARD3      LNDIWIRVWVKKGGLIVFPAGIYHRFTVDSDNYMKAMRLFVGGPVWTAYNRPHDHLPARK 180
ARD2      RNEAWIRVWVKKGGMIVLPAGIYHRFTVDSDNYIKAMRLFVGEPVWTPYNRPHDHLPARK 180
ARD4      KDDRWIRIWMQPGDLIVLPAGIYHRFTLDASNYIKLMRLFVGEPVWTPYNRPQEEHPVRK 170
           :: ***: :: *.:**:*********:*:.**:* ****** ****.****::. *.**

ARD1      EYVDNFMINASA------- 192
ARD3      AYMKKFLKVIGDRNIDASA 199
ARD2      EYIDNFVKVNEGGVIDASA 199
ARD4      KYIHGLTYKFGETVKAH-- 187
           *:. :
```

B

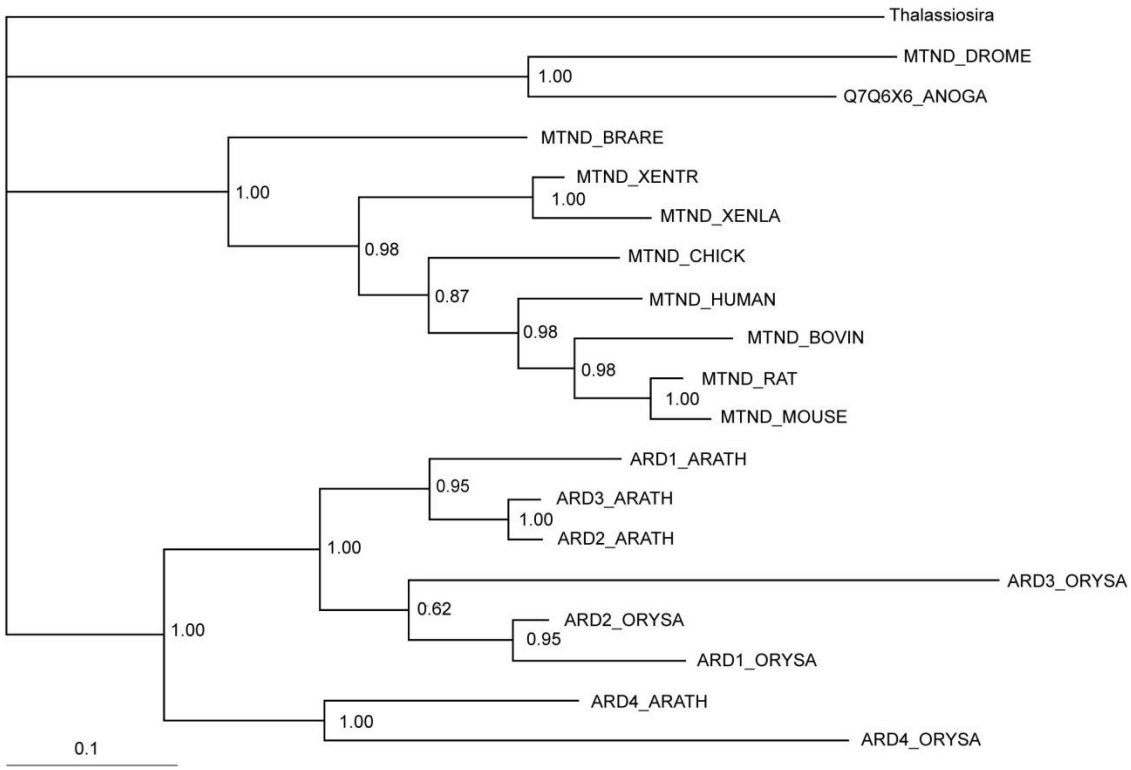


Figure 3.S3. Plants contain four ARD proteins. *A.* Multiple sequence alignment of the four *Arabidopsis* ARD proteins (red=ARD domain, underline = non-canonical NES, gray=putative metal binding residues). *(continued on next page)*

Figure 3.S3. *(continued)*
B. Phylogeny of metazoan ARD proteins generated in MrBayes. Diatom (Thalassiosira) served as an outgroup. Species indicated are as follows: ARATH, Arabidopsis; ORYSA, rice; DROME, drosophila; ANOGA, mosquito; HUMAN, human; BOVIN, bovine; RAT, rat; MOUSE, mouse; XENLA, xenopus; CHICK, chicken; BRARE, fish.

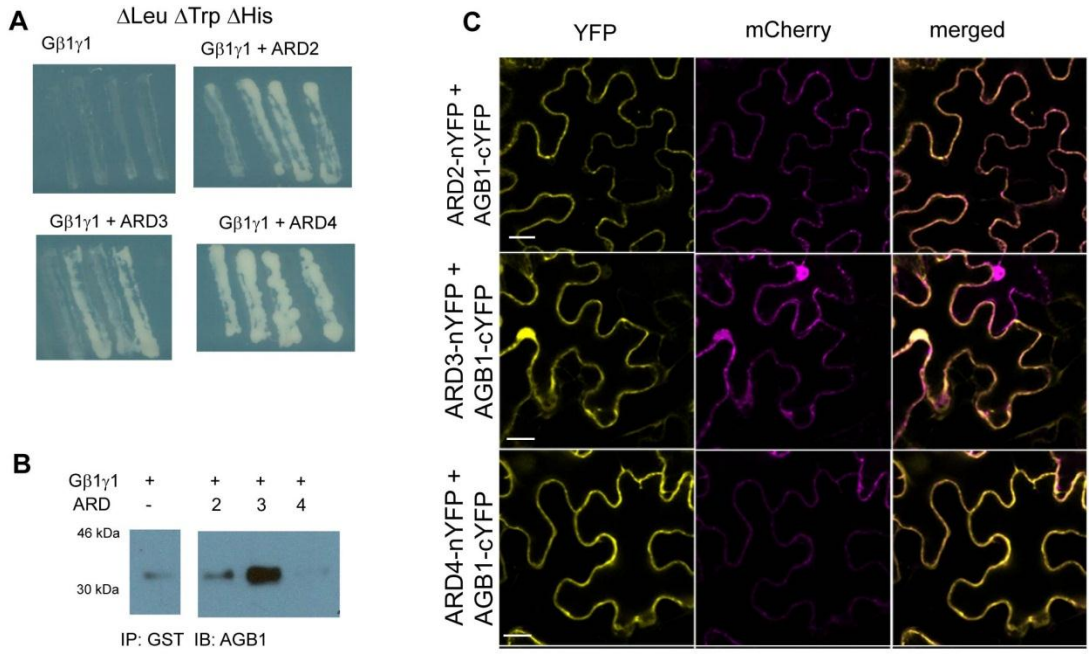


Figure 3.S4. AGB1 and ARDs interact physically. *A.* Growth of yeast strain AH109 containing the genes indicated (AtGβ1γ1 alone or AtGβ1γ1 + ARD2, ARD3, or ARD4) on yeast dropout media missing leucine, tryptophan, and histidine. This selects for a positive interaction between each of the two genes, resulting in no growth for the strain containing AtGβ1γ1 alone and positive growth for the strains containing both AtGβ1γ1 and one of the three ARDs. *B.* 6x-His-tagged Gβγ was co-immunoprecipated with ARD-GST (ARD2, ARD3, or ARD4) on a glutathione resin and detected by immunoblotting with anti-AGB1. *C.* Bifluorescence molecular complementation in tobacco leaf epidermal cells. Bar = 20 μM.

Table 3.S1. Primers used for PCR reactions. Names of primers used for RT-PCR (Figure 3.2) are denoted in parenthesis.

Name	**Purpose**	**Sequence 5'-3'**
4g14716 gtwfd1 (D)	ARD1 cDNA Gateway cloning forward	CACCATgggTgAAgCggTCAAggATg
4g14716 rv2	ARD1 reverse full-length	CTAAgCCgAggCATTgATCATgAAg
4g14716 nostpfusrv	ARD1 reverse without stop codon	AgCCgAggCATTgATCATgAAgTTA TCgACATA
ARD2 LP (C)	ARD1 Internal reverse primer (exon 3)	TAACCACgAgATTCACggATC
ARD2+3 rt-F (B)	ARD1 Internal forward primer (exon 4)	TgCCCggAAAAgCTTCCAAACTA
ARD2 rt-R (A)	ARD1 RT-PCR reverse primer	AgCCgAggCATTgATCATgAAgT
4g14710 gtwfd	ARD2 cDNA Gateway cloning forward	CACCATgggTgAAgTggTTAAggATg
4g14710 rv2	ARD2 reverse full-length	CTAggCTgACgCgTCTATgACACCA
4g14710 rvnostp	ARD2 reverse without stop codon	ggCTgACgCgTCTATgACACCACCT
2g26400 gtwfd	ARD3 cDNA Gateway cloning forward	CACCATgggTgAAgCCgCTAAggATC
2g26400 rv2	ARD3 reverse full-length	TTACgCTgAAgCgTCTATgTTACgg
2g26400 rvnostp	ARD3 reverse without stop codon	CgCTgAAgCgTCTATgTTACgggTCC
5g43850 gtwfd	ARD4 cDNA Gateway cloning forward	CACCATggCTCTCgAggCATggTTTA
5g43850 rv2	ARD4 reverse full-length	TTAATgTgCTTTAACggTTTCTCCAA ACTT
5g43850 rvnostp	ARD4 reverse without stop codon	ATgTgCTTTAACggTTTCTCCAAACT TgTAggT

References

Avila, M.A., Garcia-Trevijano, E.R., Lu, S.C., Corrales, F.J., and Mato, J.M. (2004). Methylthioadenosine. Int J Biochem Cell Biol **36,** 2125-2130.

Burstenbinder, K., Waduwara, I., Schoor, S., Moffatt, B.A., Wirtz, M., Minocha, S.C., Oppermann, Y., Bouchereau, A., Hell, R., and Sauter, M. (2010). Inhibition of 5'-methylthioadenosine metabolism in the Yang cycle alters polyamine levels, and impairs seedling growth and reproduction in Arabidopsis. Plant J **62,** 977-988.

Chai, S.C., Ju, T., Dang, M., Goldsmith, R.B., Maroney, M.J., and Pochapsky, T.C. (2008). Characterization of metal binding in the active sites of acireductone dioxygenase isoforms from Klebsiella ATCC 8724. Biochemistry **47,** 2428-2438.

Dai, Y., Wensink, P.C., and Abeles, R.H. (1999). One protein, two enzymes. J Biol Chem **274,** 1193-1195.

Dai, Y., Pochapsky, T.C., and Abeles, R.H. (2001). Mechanistic studies of two dioxygenases in the methionine salvage pathway of Klebsiella pneumoniae. Biochemistry **40,** 6379-6387.

Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.S., Iyengar, R., Miller, R.J., Jan, L.Y., Lefkowitz, R.J., and Hamm, H.E. (1998). Molecular basis for interactions of G protein betagamma subunits with effectors. Science **280,** 1271-1274.

Friedman, E.J., Temple, B.R., Hicks, S.N., Sondek, J., Jones, C.D., and Jones, A.M. (2009). Prediction of protein-protein interfaces on G-protein beta subunits reveals a novel phospholipase C beta2 binding domain. J Mol Biol **392,** 1044-1054.

Gotoh, I., Uekita, T., and Seiki, M. (2007). Regulated nucleo-cytoplasmic shuttling of human aci-reductone dioxygenase (hADI1) and its potential role in mRNA processing. Genes Cells **12,** 105-117.

Grigston, J.C., Osuna, D., Scheible, W.R., Liu, C., Stitt, M., and Jones, A.M. (2008). D-Glucose sensing by a plasma membrane regulator of G signaling protein, AtRGS1. FEBS Lett **582,** 3577-3584.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., and Bollback, J.P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. Science **294,** 2310-2314.

Jeffery, C.J. (2003). Moonlighting proteins: old proteins learning new tricks. Trends Genet **19,** 415-417.

Ju, T., Goldsmith, R.B., Chai, S.C., Maroney, M.J., Pochapsky, S.S., and Pochapsky, T.C. (2006). One protein, two enzymes revisited: a structural entropy switch interconverts the two isoforms of acireductone dioxygenase. J Mol Biol **363,** 823-834.

Kim, J.H., Kim, H.S., Lee, Y.H., Kim, Y.S., Oh, H.W., Joung, H., Chae, S.K., Suh, K.H., and Jeon, J.H. (2008). Polyamine biosynthesis regulated by StARD expression plays an important role in potato wound periderm formation. Plant Cell Physiol **49,** 1627-1632.

Lin, T., He, X., Yang, L., Shou, H., and Wu, P. (2005). Identification and characterization of a novel water-deficit-suppressed gene OsARD encoding an aci-reductone-dioxygenase-like protein in rice. Gene **360,** 27-34.

Llorente, F., Alonso-Blanco, C., Sanchez-Rodriguez, C., Jorda, L., and Molina, A. (2005). ERECTA receptor-like kinase and heterotrimeric G protein from Arabidopsis are required for resistance to the necrotrophic fungus Plectosphaerella cucumerina. Plant J **43,** 165-180.

Mudgil, Y., Uhrig, J.F., Zhou, J., Temple, B., Jiang, K., and Jones, A.M. (2009). Arabidopsis N-MYC DOWNREGULATED-LIKE1, a positive regulator of auxin transport in a G protein-mediated pathway. Plant Cell **21,** 3591-3609.

Oram, S.W., Ai, J., Pagani, G.M., Hitchens, M.R., Stern, J.A., Eggener, S., Pins, M., Xiao, W., Cai, X., Haleem, R., Jiang, F., Pochapsky, T.C., Hedstrom, L., and Wang, Z. (2007). Expression and function of the human androgen-responsive gene ADI1 in prostate cancer. Neoplasia **9,** 643-651.

Oredsson, S.M. (2003). Polyamine dependence of normal cell-cycle progression. Biochem Soc Trans **31,** 366-370.

Pirkov, I., Norbeck, J., Gustafsson, L., and Albers, E. (2008). A complete inventory of all enzymes in the eukaryotic methionine salvage pathway. Febs J **275,** 4111-4120.

Pochapsky, T.C., Pochapsky, S.S., Ju, T., Hoefler, C., and Liang, J. (2006). A refined model for the structure of acireductone dioxygenase from Klebsiella ATCC 8724 incorporating residual dipolar couplings. J Biomol NMR **34,** 117-127.

Pochapsky, T.C., Pochapsky, S.S., Ju, T., Mo, H., Al-Mjeni, F., and Maroney, M.J. (2002). Modeling and experiment yields the structure of acireductone dioxygenase from Klebsiella pneumoniae. Nat Struct Biol **9,** 966-972.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol **53,** 247-259.

Sauter, M., Lorbiecke, R., Ouyang, B., Pochapsky, T.C., and Rzewuski, G. (2005). The immediate-early ethylene response gene OsARD1 encodes an acireductone dioxygenase involved in recycling of the ethylene precursor S-adenosylmethionine. Plant J **44,** 718-729.

Smrcka, A.V. (2008). G protein betagamma subunits: central mediators of G protein-coupled receptor signaling. Cell Mol Life Sci **65,** 2191-2214.

Sundin, A.M., and Larsson, J.E. (2002). Rapid and sensitive method for the analysis of carbon monoxide in blood using gas chromatography with flame ionisation detection. J Chromatogr B Analyt Technol Biomed Life Sci **766,** 115-121.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25,** 4876-4882.

Trusov, Y., Sewelam, N., Rookes, J.E., Kunkel, M., Nowak, E., Schenk, P.M., and Botella, J.R. (2008). Heterotrimeric G proteins-mediated resistance to necrotrophic pathogens includes mechanisms independent of salicylic acid-, jasmonic acid/ethylene- and abscisic acid-mediated defense signaling. Plant J.

Tsuchisaka, A., Yu, G., Jin, H., Alonso, J.M., Ecker, J.R., Zhang, X., Gao, S., and Theologis, A. (2009). A combinatorial interplay among the 1-aminocyclopropane-1-carboxylate isoforms regulates ethylene biosynthesis in Arabidopsis thaliana. Genetics **183,** 979-1003.

Ullah, H., Chen, J.G., Young, J.C., Im, K.H., Sussman, M.R., and Jones, A.M. (2001). Modulation of cell proliferation by heterotrimeric G protein in Arabidopsis. Science **292,** 2066-2069.

Ullah, H., Chen, J.G., Temple, B., Boyes, D.C., Alonso, J.M., Davis, K.R., Ecker, J.R., and Jones, A.M. (2003). The beta-subunit of the Arabidopsis G protein negatively regulates auxin-induced cell division and affects multiple developmental processes. Plant Cell **15,** 393-409.

Vogel, J.P., Woeste, K.E., Theologis, A., and Kieber, J.J. (1998). Recessive and dominant mutations in the ethylene biosynthetic gene ACS5 of Arabidopsis confer cytokinin insensitivity and ethylene overproduction, respectively. Proc Natl Acad Sci U S A **95,** 4766-4771.

Wang, H.X., Weerasinghe, R.R., Perdue, T.D., Cakmakci, N.G., Taylor, J.P., Marzluff, W.F., and Jones, A.M. (2006). A Golgi-localized hexose transporter is involved in heterotrimeric G protein-mediated early development in Arabidopsis. Mol Biol Cell **17,** 4257-4269.

Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.Y., Xia, Y., Dixon, R.A., Harrison, M.J., Lamb, C.J., Yanofsky, M.F., and Chory, J. (2000). Activation tagging in Arabidopsis. Plant Physiol **122,** 1003-1013.

Woeste, K., and Kieber, J.J. (1998). The molecular basis of ethylene signalling in Arabidopsis. Philos Trans R Soc Lond B Biol Sci **353,** 1431-1438.

Wray, J.W., and Abeles, R.H. (1995). The methionine salvage pathway in Klebsiella pneumoniae and rat liver. Identification and characterization of two novel dioxygenases. J Biol Chem **270,** 3147-3153.

Xu, Q., Schwarzenbacher, R., Krishna, S.S., McMullan, D., Agarwalla, S., Quijano, K., Abdubek, P., Ambing, E., Axelrod, H., Biorac, T., Canaves, J.M., Chiu, H.J., Elsliger, M.A., Grittini, C., Grzechnik, S.K., DiDonato, M., Hale, J., Hampton, E., Han, G.W., Haugen, J., Hornsby, M., Jaroszewski, L., Klock, H.E., Knuth, M.W., Koesema, E., Kreusch, A., Kuhn, P., Miller, M.D., Moy, K., Nigoghossian, E., Paulsen, J., Reyes, R., Rife, C., Spraggon, G., Stevens, R.C., van den Bedem, H., Velasquez, J., White, A., Wolf, G., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., and Wilson, I.A. (2006). Crystal structure of acireductone dioxygenase (ARD) from Mus musculus at 2.06 angstrom resolution. Proteins **64,** 808-813.

Yoo, S.D., Cho, Y.H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc **2,** 1565-1572.

Zhang, Y., Heinsen, M.H., Kostic, M., Pagani, G.M., Riera, T.V., Perovic, I., Hedstrom, L., Snider, B.B., and Pochapsky, T.C. (2004). Analogs of 1-phosphonooxy-2,2-dihydroxy-3-oxo-5-(methylthio)pentane, an acyclic intermediate in the methionine salvage pathway: a new preparation and characterization of activity with E1 enolase/phosphatase from Klebsiella oxytoca. Bioorg Med Chem **12,** 3847-3855.

CHAPTER 4

Discussion and Future Directions

Summary

Cell signaling through the heterotrimeric G protein complex is universally employed in eukaryotes, and it mediates such critical processes as growth, proliferation, and stress responses (Jones and Assmann, 2004). GPCRs are the cellular targets for approximately 40% of all prescription pharmaceuticals marketed today (Filmore, 2004). Therefore, it is critical that we understand the function of both G protein components and their downstream targets. My work specifically involved the identification and characterization of interactions between the Gβ subunit of the heterotrimeric G protein and its downstream effectors.

We first took an evolutionary approach to identify novel regions of non-overlapping protein-protein interactions on the Gβ protein surface, as described in Chapter 2. We utilized this type of approach for several reasons: 1) Gβ proteins have duplicated and diverged over time, 2) Known proteins that interact with Gβs have arisen throughout evolutionary history, 3) Structural data for Gβ proteins in complex with interacting proteins is limited, and 4) The high utilization of the Gα-Gβ binding interface by effectors has masked the identification of additional binding interfaces. As a result of our work, we were able to confirm the classification of Gβ subunits into two classes (Gβ1-like and Gβ5-like). We identified a novel interaction between Gβ and the interacting protein phosducin in plants, we identified and experimentally confirmed a novel binding interface between mammalian Gβ1 and the effector phospholipase C-β2, and we identified six additional putative interfaces with no previously-known function (Table 2.1).

The plant *Arabidopsis thaliana* is an ideal system in which to study G protein signaling due to its genetic simplicity yet homology to mammalian systems. Therefore, we took a second approach to identify Gβ interacting proteins in *Arabidopsis*, as described in Chapter 3. We used activation tagging to identify genetic suppressors of *agb1-2*, and this screen identified *ARD1*. We were able to show that ARD1 and AGB1 interact genetically and physically, and we further showed that AGB1 directly stimulates ARD1 enzymatic activity. We utilized the starting point for the evolutionary analysis described in Chapter 2 to identify and confirm sites of protein-protein interaction between ARD1 and AGB1. These sites allowed us to predict and test a mechanism whereby AGB1 stimulates ARD1 enzymatic activity by "unlocking" the molecule at W166 and displacing the C-terminal helix.

By taking this combination of approaches, we were able to answer structure-function questions in the absence of an AGB1 crystal structure (either alone or in complex with ARD1). Additionally, we were able to utilize the knowledge gained from our evolutionary analysis to make predictions regarding the interaction between AGB1 and the novel effector ARD1. However, while these results answer several questions, they raise several more.

Evolution of ARDs in plants

As mentioned in the Discussion of Chapter 3, organisms from bacteria through mammals each contain one ARD protein. However, plant species contain multiple copies of ARDs; for example, rice and *Arabidopsis* each contain four ARDs (Figure 3.S3B). Of these four proteins, three are highly similar (80-95% identity), while the fourth has diverged (75% identity between ARD1 and ARD4). It is intriguing that multiple ARDs

evolved and were maintained solely in plants. To identify both overlapping and non-overlapping functions of the four ARDs, we first need to create higher-order mutants within the family. An artificial microRNA (amiRNA) construct was created (Schwab et al., 2006) that targets ARD1, ARD2, and ARD3 and a second construct that targets only ARD4. Both wildtype and *agb1-2* plants were transformed with amiRNA ARD123, amiRNA ARD4, or the combination of amiRNA ARD123 + amiRNA ARD4. However, in these plants, the *ARD* genes were knocked down but not knocked out (Fig 4.1). To date, no altered phenotypes deriving from these reduced transcripts have been identified, suggesting that we may require null alleles of each gene. Three null alleles of *ard1* were identified, and we are in the process of confirming multiple null alleles of *ard4*. Even if null mutations in all four ARDs are identified, a quadruple null mutant is unlikely to ever be generated by crossing due to the short mapping distance between ARD1 and ARD2 (adjacent genes on chromosome four). However, by creating a null mutant of *ard1 ard2 ard4* that is transformed with amiRNA targeting ARD2, we can create a plant that lacks ARD1, ARD3, ARD4, and contains a reduced level of ARD2. It is possible that this scenario would allow us to see the *in planta* effects of mutating ARDs.

ARD1 stimulation by AGB1

We showed in Chapter 4 that AGB1 can stimulate ARD1 enzymatic activity by "unlocking" W166 and displacing the C-terminal helix (Figure 3.5). Previous structural studies have shown that the region surrounding the active site of ARD1 exists in multiple states based on the organism of purification and the metal bound in the active site (Ju et al., 2006; Pochapsky et al., 2006; Xu et al., 2006); the active site is the most "closed" in mADI1, the most "open" in bacterial Fe-ARD, and is intermediate in bacterial Ni-ARD.

In order to determine the function of W166, we are currently attempting to crystallize ARD1 wildtype and ARD1 W166A. This would allow for the comparison between the two structures and the identification of structural differences between the basal and stimulated forms of ARD1 in order to determine how ARD1 is regulated structurally. However, crystallography may be impeded by the fact that wildtype ARD1 is a highly unstable enzyme (as discussed in Chapter 3). We are not yet certain of the cause behind the dramatic drop in enzymatic activity over time, but we cannot rule out that it is in part due to structural instability. Our collaborators Tom Pochapsky and Iva Perovic have demonstrated that the mutant W166A is stable over time (Perovic and Pochapsky, unpublished data). Therefore, in the event that wildtype ARD1 cannot be crystallized, the stable ARD1 W166A may be used as a tool for generating a crystal structure. Even in the absence of a wildtype ARD1 structure, we may still make comparisons between ARD1 W166A, existing mammalian and bacterial crystal structures, and the homology model of ARD1 (Figure 3.S1B).

Function of specific AGB1 residues in plants

Although we have shown that mutations in AGB1 can affect the ability of AGB1 to stimulate ARD1 enzymatic activity, we also wanted to know what effects these mutations conferred on plants. *agb1*-2 null mutant plants display a wide array of phenotypes (discussed in Chapter 1) that are likely mediated by several downstream effectors that may bind to different AGB1 surfaces. In order to dissect the specific phenotypic contributions of several different surfaces of AGB1, we expressed each AGB1 mutation under the control of the strong 35S promoter in *agb1-2.* I am currently mentoring Arwen Frick-Cheng (an undergraduate researcher) and Kun Jiang (a

postdoctoral researcher) as they assay each published *agb1-2* phenotype in this set of mutant plants. Together, we are investigating the contribution of each surface on the two-day-old etiolated hypocotyl length and hook angle, lateral root production, cotyledon shape, petiole length, leaf shape, silique shape, and glucose response. Further, we have initiated collaborations with two groups that are investigating the resistance of each mutant to fungal pathogens. The preliminary results of this project indicate that each AGB1 mutation fully rescues at least one *agb1-2* phenotype, and each mutated AGB1 surface confers a specific combination of *agb1-2*-like phenotypes. These data will allow us to further dissect the AGB1 signaling pathway and to ascribe functionality to several portions of the AGB1 protein surface.

Proteins sharing binding sites with ARD1

Although we have shown that three of the four mutations created in AGB1 reduced ARD1 stimulation by AGB1 (Figure 3.4), preliminary data suggest that these sites confer additional AGB1 function (Frick-Cheng and Jiang, unpublished data, see previous section). Recently, our lab has completed a large-scale G protein interactome collaborative project. Briefly, known G protein components were used as baits in a Y2H screen against multiple genomic libraries. Positive results were confirmed via one-on-one Y2H screening, and a subset of these proteins was tested for *in vivo* interaction via BiFC. As a result of this project, we have compiled a large list of proteins that physically interact with AGB1. In order to determine whether any of these proteins bind the residues that were mutated (see previous section), the four mutant AGB1 constructs could be cloned into Y2H vectors and tested 1:1 for interaction with known binding partners identified in the interactome. Any protein that fails to interact with an AGB1 mutant may

require those mutated residues for binding to AGB1. As a result, additional downstream targets that confer the phenotypes discussed in the previous section may be identified.

Additional functions of ARDs

The ARDs studied to date perform the methionine salvage step that converts aci-reductone into the keto-acid methionine precursor. However, several additional functions for ARD proteins have been identified in mammals, and some of these do not rely on the enzymatic activity of ARD. For example, human ARD is localized in the cytoplasm, nucleus, and at the plasma membrane (Uekita et al., 2004); it regulates mRNA splicing (Gotoh et al., 2007) and stimulates apoptosis (Oram et al., 2004; Uekita et al., 2004) independently of its enzymatic function. Through the G protein interactome project (discussed above), we identified an interaction between ARD1 and the transcription factor MYC2 (also known as JIN1). Interestingly, a genetic relationship between *MYC2* and *AGB1* had been previously identified. Together, *AGB1* and *MYC2* modulate the resistance to the pathogen *F. oxysporum* (Trusov et al., 2008). We confirmed that the ARD1-MYC2 physical interaction occurs *in vivo* in the nucleus with BiFC (Figure 4.2), suggesting a nuclear role for ARD1 function in plants. Our collaborators Yuri Trusov and Jimmy Botella have further shown that *ard4* mutant plants (but not *ard1*) are resistant to *F. oxysporum* infection, while *agb1-2* is hypersensitive. As a result, Kun Jiang, a post-doc in our lab, is currently investigating the contribution of ARDs and AGB1 to disease resistance. It is possible that AGB1 interacts with ARD1 and/or ARD4 to mediate this resistance via a physical nuclear interaction between ARDs and MYC2.

G protein regulation of mammalian ARD

Although the plant-specific findings regarding ARD function and regulation are important for the plant community, we also strive to use *Arabidopsis* as a model system for mammalian G protein signaling. ARD has previously been studied in mammalian systems and has been shown to play roles in cancer growth and cell death, but its regulation by G proteins has never been proposed. Therefore, it is critical that we determine whether mammalian ARD is also regulated by Gβγ. Our preliminary data suggest that this may be true; mouse ARD was able to interact physically with *Arabidopsis* Gβγ in a Y3H screen (Figure 4.3). This interaction needs to be confirmed in higher eukaryotes. Because mammals contain up to 60 Gβγ dimer combinations, *C. elegans* (which contain members of both classes of Gβ, but many fewer Gβ and Gγ genes than mammals) would be a more feasible model system in which to confirm this hypothesis.

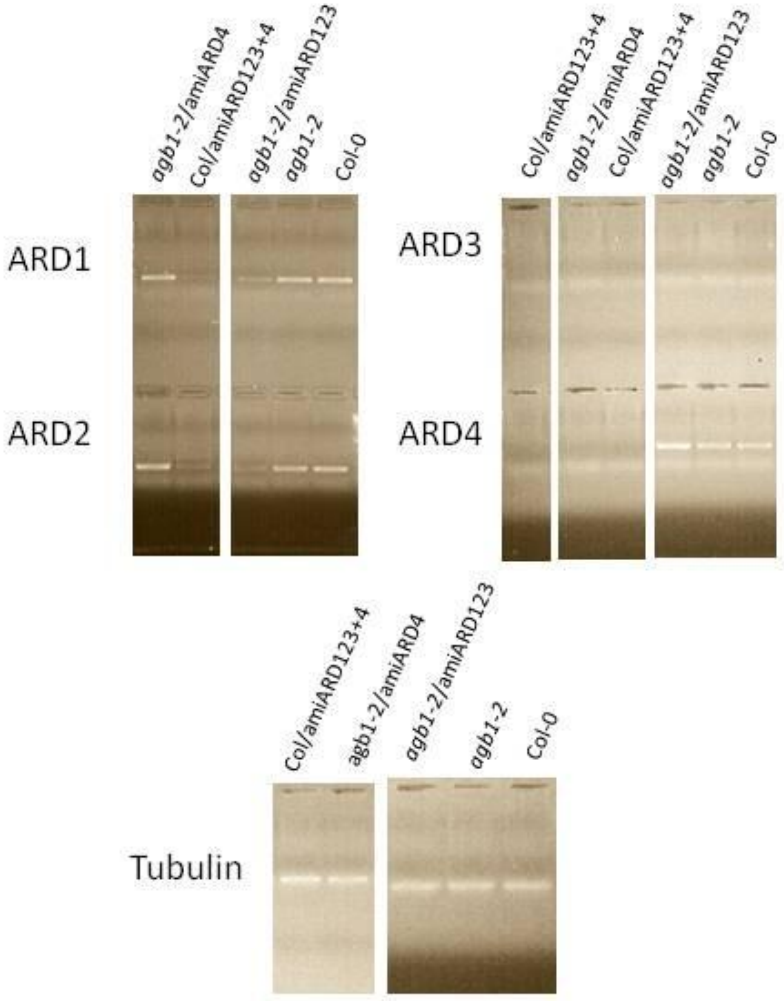


Figure 4.1. amiRNA constructs reduce the transcript levels of ARD mRNA. Transcript levels of *ARD1*, *ARD2*, *ARD3*, and *ARD4* mRNA as determined by RT-PCR. RT-PCR was performed using the primers indicated. *Tubulin4* was used as a reference transcript. PCR was performed for 25 cycles, and PCR products were run on one gel in non-adjacent lanes. Note that *ARD3* levels were not detectable even in *agb1-2* and Col-0 plants.

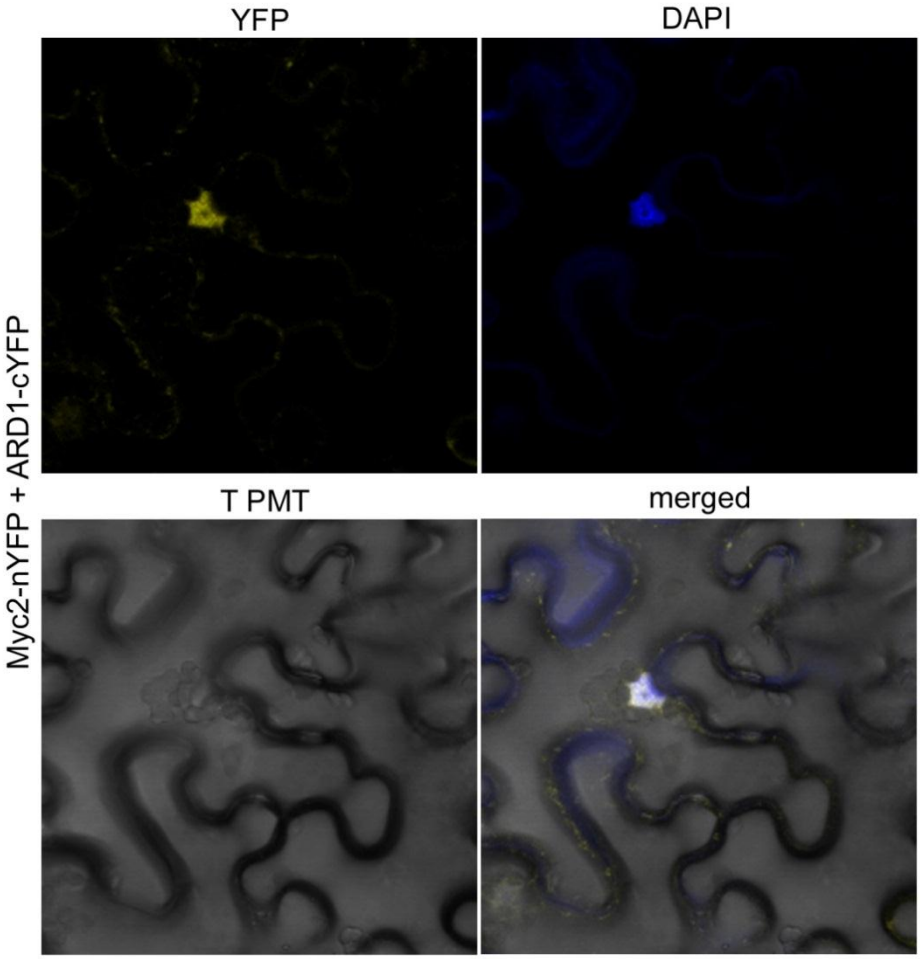


Figure 4.2. ARD1 physically interacts with Myc2 in the nucleus. Bifluorescence molecular complementation in *N. benthamiana* leaf epidermal cells confirms that ARD1 interacts with Myc2 in the nucleus as indicated by the reconstitution of the YFP signal and its co-localization with DAPI staining. T PMT channel shows the cell borders using transmitted light.

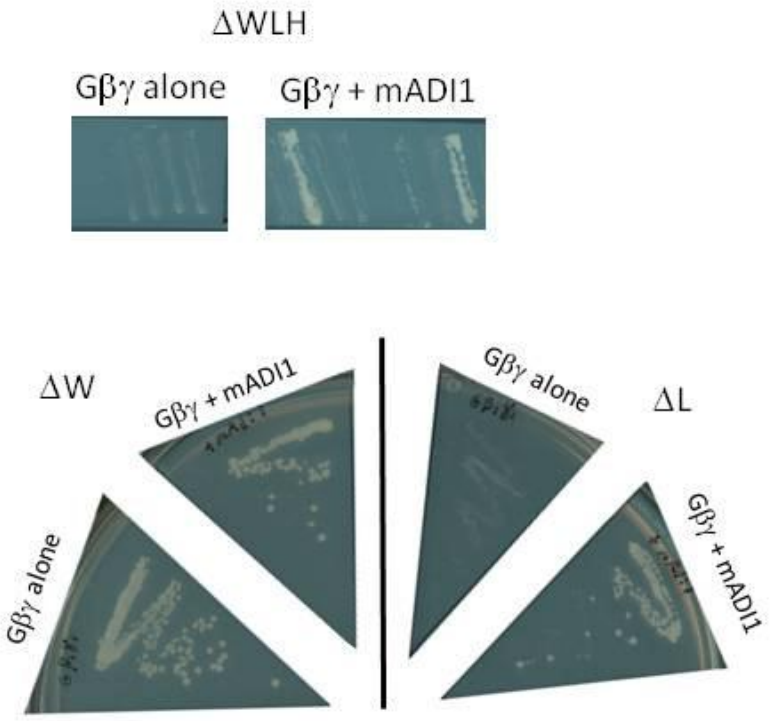


Figure 4.3. The AGB1-AGG1 heterodimer physically interacts with mouse ARD. Growth of yeast strain AH109 containing the proteins indicated (AGB1-AGG1 (AtGβγ) alone or AtGβγ + mADI1) on yeast dropout media missing leucine, tryptophan, and histidine. This selects for a positive interaction between each of the two genes, resulting in no growth for the strain containing AtGβγ alone and positive growth for the strain containing both AtGβγ and ARD1. Four to six independent colonies were restreaked on the triple dropout media to confirm the interaction. Positive growth on media lacking leucine or tryptophan confirms the presence of mADI1 or AtGβγ, respectively.

References

Filmore, D. (2004). It's a GPCR world. Modern Drug Discovery **7,** 24-28.

Gotoh, I., Uekita, T., and Seiki, M. (2007). Regulated nucleo-cytoplasmic shuttling of human aci-reductone dioxygenase (hADI1) and its potential role in mRNA processing. Genes Cells **12,** 105-117.

Jones, A.M., and Assmann, S.M. (2004). Plants: the latest model system for G-protein research. EMBO Rep **5,** 572-578.

Ju, T., Goldsmith, R.B., Chai, S.C., Maroney, M.J., Pochapsky, S.S., and Pochapsky, T.C. (2006). One protein, two enzymes revisited: a structural entropy switch interconverts the two isoforms of acireductone dioxygenase. J Mol Biol **363,** 823-834.

Oram, S., Jiang, F., Cai, X., Haleem, R., Dincer, Z., and Wang, Z. (2004). Identification and characterization of an androgen-responsive gene encoding an aci-reductone dioxygenase-like protein in the rat prostate. Endocrinology **145,** 1933-1942.

Pochapsky, T.C., Pochapsky, S.S., Ju, T., Hoefler, C., and Liang, J. (2006). A refined model for the structure of acireductone dioxygenase from Klebsiella ATCC 8724 incorporating residual dipolar couplings. J Biomol NMR **34,** 117-127.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D. (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. Plant Cell **18,** 1121-1133.

Trusov, Y., Sewelam, N., Rookes, J.E., Kunkel, M., Nowak, E., Schenk, P.M., and Botella, J.R. (2008). Heterotrimeric G proteins-mediated resistance to necrotrophic pathogens includes mechanisms independent of salicylic acid-, jasmonic acid/ethylene- and abscisic acid-mediated defense signaling. Plant J.

Uekita, T., Gotoh, I., Kinoshita, T., Itoh, Y., Sato, H., Shiomi, T., Okada, Y., and Seiki, M. (2004). Membrane-type 1 matrix metalloproteinase cytoplasmic tail-binding protein-1 is a new member of the Cupin superfamily. A possible multifunctional protein acting as an invasion suppressor down-regulated in tumors. J Biol Chem **279,** 12734-12743.

Xu, Q., Schwarzenbacher, R., Krishna, S.S., McMullan, D., Agarwalla, S., Quijano, K., Abdubek, P., Ambing, E., Axelrod, H., Biorac, T., Canaves, J.M., Chiu, H.J., Elsliger, M.A., Grittini, C., Grzechnik, S.K., DiDonato, M., Hale, J., Hampton, E., Han, G.W., Haugen, J., Hornsby, M., Jaroszewski, L., Klock, H.E., Knuth, M.W., Koesema, E., Kreusch, A., Kuhn, P., Miller, M.D., Moy,

K., Nigoghossian, E., Paulsen, J., Reyes, R., Rife, C., Spraggon, G., Stevens, R.C., van den Bedem, H., Velasquez, J., White, A., Wolf, G., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., and Wilson, I.A. (2006). Crystal structure of acireductone dioxygenase (ARD) from Mus musculus at 2.06 angstrom resolution. Proteins **64,** 808-813.